

Open Research Online

The Open University's repository of research publications and other research outputs

Family and tumour studies in breast and oesophageal cancer

Thesis

How to cite:

Kellsell, David Peter (1997). Family and tumour studies in breast and oesophageal cancer. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 1996 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000d525>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Family and Tumour Studies in Breast and Oesophageal Cancer

David Peter Kelsell

**A thesis submitted in fulfilment of the requirements
of the Open University for the degree of
Doctor of Philosophy.**

September 1996

**The Imperial Cancer Research Fund
44 Lincoln's Inn Field
London WC2A 3PX**

Author's number : P9275010

Date of submission : September 1996

Date of award : 9th January 1997

Acknowledgements

A big thank-you to my two supervisors, Nigel Spurr and Tim Bishop, for their guidance and discussions in the initiation and the completion of this work. I would like to acknowledge the patients and family members who provided the material which is the basis of this study. Thanks go to all the collaborators on the various projects encompassed in this study, in particular, the pedigree workers who ascertained the breast cancer families, Diana Barnes who provided the breast tumour material and Irene Leigh's group who provided the material for the NEPPK study. I would like to thank the members of the Human Genetic Resources laboratory for all their help, in particular, Ian Gray who provided useful discussions and cast-iron methods. Special thanks go to Rosemary Kelsell and Ian Gray for their invaluable contributions in the preparation of this manuscript. I would also like to acknowledge the following people: Tim Bishop and Steve Bryant for linkage analyses; Bim Laguda and Lesley Rooke for DNA extraction; Tomas Lindahl for his support; John Nicholson for photography; and Iain Goldsmith's group for the extensive supply of oligonucleotides. Finally, I would like to thank my wife Rosemary, colleagues and members of my family for their interest and support during the course of this study.

Abstract

This study focussed on two areas in the field of cancer susceptibility. The initial area was the genetic analysis of a recently mapped breast cancer susceptibility locus, BRCA1, in a number of breast and breast-ovarian cancer families. In the largest of the ICRF families studied (BOV3), linkage to the long arm of chromosome 17 was confirmed and a number of recombinants were identified. One such cross-over event enabled the reduction of the interval harbouring BRCA1 to a region estimated to be between 1-1.5 Mb. During the course of this study, a second gene for breast/ovarian cancer predisposition, BRCA2, had been assigned to a 6 cM region at 13q12-13. Towards the identification of this gene, a YAC contig was constructed spanning the published minimal genetic interval for BRCA2. This contig provided a framework for the identification of BRCA2. Allele loss studies were also performed and indicated that BRCA1 acts as a tumour suppressor. Analysis of familial and sporadic infiltrating ductal grade 3 breast carcinomas revealed a pattern of combined loss or retention of BRCA1 and BRCA2. This supports a role for both genes in the development of this tumour type. The other area of study was the genetic analysis of a group of autosomal dominant skin diseases, termed the non-epidermolytic palmoplantar keratodermas. This study demonstrated genetic heterogeneity between three forms of NEPPK: diffuse, punctate and focal. Genetic heterogeneity was also established between families presenting with clinically similar forms of focal NEPPK. Mutations in the type I keratin on 17q12-21, KRT16, were identified as the genetic basis of focal NEPPK in a pedigree without associated susceptibility to oesophageal cancer. In the pedigrees with a striking association between focal NEPPK and oesophageal cancer susceptibility, the region harbouring this disease locus (TOC) was refined to a 1cM region on 17q24-25.

TABLE OF CONTENTS

Page Number

Title	1
Acknowledgements	2
Abstract	3
Table of Contents	4
Abbreviations	10

CHAPTER 1: INTRODUCTION 12

1.1. Breast Cancer	12
1.2. Risk Factors	12
1.3. Segregation Analysis	16
1.4. Linkage Analysis	17
1.5. Genetic Markers	21
1.6. Cancer is a multi-step genetic disease	27
1.7. Cancer Genes	29
1.8. Tumour Suppressor Genes	29
1.9. Oncogenes	37
1.10. Objectives of study	39

CHAPTER 2: MATERIALS and METHODS 41

2.1. Family Ascertainment	41
2.2. Source of DNA	42
2.3. Extraction of Genomic DNA	45
2.4. Restriction Enzyme Analysis	49
2.5. Southern Blotting	51
2.6. DNA Amplification by the Polymerase Chain Reaction	54
2.7. Detection of alleles at microsatellite marker loci	57
2.8. Assessment of allele loss	60
2.9. NME2 cloning	60
2.10. Sequence Analysis of DNA	61
2.11. YAC identification	65
2.12. BAC library screening by PCR amplification	67
2.13. Linkage analysis	67
2.14. Estimation of allele frequencies	69

<u>CHAPTER 3: GENETIC MAPPING OF BRCA1</u>	70
3.1. Introduction	70
3.2. Linkage analysis of 16 breast and breast-ovarian cancer families	70
3.3. Breast Cancer Linkage Consortium Analysis	79
3.4. Re-evaluation of ICRF mapping study	79
3.5. Further genetic analysis of BRCA1 in BOV3	81
3.6. Candidate gene analysis for BRCA1	88
3.7. NME 1 and 2 (NM23)	89
3.8. Cloning of NME2	89
3.9. Localisation of NME2 by Southern hybridisation	90
3.10. Localisation of NME2 to 17q21-22 by PCR amplification	93
3.11. Genetic exclusion of other BRCA1 candidate genes	96
3.12. DNA mutation analysis of EDH17B1 and EDH17B2	96
3.13. Summary of Chapter 3	99
3.12. The cloning of the BRCA1 gene	102
<u>CHAPTER 4: PHYSICAL MAP OF THE BRCA2 REGION</u>	104
4.1. Introduction	104
4.2. Physical map of the BRCA2 region	109
4.3. Assignment of BRCA2 candidate genes	116
4.4. Summary	119
4.5. Identification of BRCA2	120
<u>CHAPTER 5: ALLELE LOSS AT BRCA1 AND BRCA2</u>	121
5.1. Introduction	121
5.2. Is BRCA1 a tumour suppressor gene?	121
5.3. Allele loss at BRCA1 in familial breast tumours	129
5.4. BRCA1 allele loss in sporadic breast carcinomas	148
5.5. BRCA2 allele loss in sporadic breast carcinomas	149
5.6. Allele loss of BRCA1/ BRCA2 in grade 3 breast carcinomas	151
5.7. Summary	157

CHAPTER 6: PALMOPLANTAR KERATODERMA AND

CANCER SUSCEPTIBILITY 158

6.1. Concept	158
6.2. Keratin diseases	158
6.3. Palmoplantar keratoderma	159
6.4. Aim of study	160
6.5. Diffuse NEPPK	161
6.6. Focal NEPPK without associated Oesophageal Cancer	167
6.7. Focal NEPPK and Oesophageal Cancer	171
6.8. Linkage analysis in pedigree 2216	171
6.9. Fine mapping of the TOC locus	177
6.10. Ancestral TOC mutation?	181
6.11. Does TOC act as a tumour suppressor gene?	182
6.12. Candidate genes for TOC?	186
6.13. Summary of TOC analysis	197
6.14. Punctate NEPPK	197
6.15. Summary of Chapter 6	201

CHAPTER 7. DISCUSSION 202

7.1. Introduction	202
7.2. Breast-ovarian cancer	202
7.2.1. Linkage study	202
7.2.2. Role of BRCA1 and BRCA2 in sporadic breast cancer	205
7.2.3. Function of BRCA1	209
7.3. NEPPK and cancer susceptibility	210
7.3.1. Genetic heterogeneity	210
7.3.2. Focal NEPPK and Cancer susceptibility	211
7.3.3. Punctate PPK and Cancer	214
7.4. Concluding remarks	214

REFERENCES 216

FIGURES

Figure 1.1. The affect of allele number and frequency in linkage analysis	24
Figure 1.2. Multistep model for colorectal tumourigenesis	28
Figure 1.3. Knudson's two hit model for retinoblastoma	31
Figure 1.4. Allele loss to detect tumour suppressor genes	33
Figure 3.1. Genetic map of DNA marker loci mapping to the long arm of chromosome 17	71
Figure 3.2. Haplotype construction for 17q DNA marker loci in BOV3	75
Figure 3.3. Haplotype construction for 17q DNA marker loci in families: (a)BOV2 and (b) BC19	78
Figure 3.4. Haplotype construction of 17q DNA marker loci mapping between D17S250 and D17S293 in the BOV3 family	83
Figure 3.5. Mapping the recombination event in individual 3684 from BOV3	85
Figure 3.6 Reduction of the minimal region for BRCA1	86
Figure 3.7. Heterozygosity at D17S800 in individual 5367	87
Figure 3.8. Southern hybridisation of NME2 to 17q21-22	92
Figure 3.9. Localisation of NME2 to 17q21-22 by PCR amplification	95
Figure 3.10. Exclusion of candidate genes for BRCA1	97
Figure 3.11. Sequencing of EDH17B1 and EDH17B2	101
Figure 3.12. Map location of BRCA1	102
Figure 4.1. Three breast and breast-ovarian cancer families not linked to BRCA1	106
Figure 4.2. ICRF26, a breast cancer family not linked to BRCA1	108
Figure 4.3. YAC characterisation of the BRCA2 minimal region	113
Figure 4.4. Graphic representation of the minimal tiling path of YACs spanning BRCA2 on 13q12-13	115
Figure 4.5. D13S178E, a candidate gene for BRCA2?	118
Figure 5.1. Allele loss on chromosome 17 in BOV3 breast tumours	123
Figure 5.2. Examples of chromosome 17 allele status in breast carcinomas from BOV3	124
Figure 5.3. Loss of the wildtype BRCA1 chromosome in breast tumours from the BOV3 family	127
Figure 5.4. Haplotype and allele loss analysis of BC5	132
Figure 5.5. Haplotyping and allele loss analysis of BOV4	135
Figure 5.6. Allele loss at the THRA1 locus in ovarian tumour DNA of individual 8190 from BOV4	136

Figure 5.7. BRCA1 haplotype and allele loss analysis in three ICRF pedigrees	140
Figure 5.8. ICRF510 BRCA1 haplotype and tumour DNA analysis	146
Figure 5.9. Allele loss at the THRA1 locus in breast tumour DNA of individual 12259 from ICRF510	147
Figure 5.10. Pattern of allele loss on 17q in DNA from a “sporadic” breast tumour 1133	150
Figure 5.11. Pattern of allele loss in tumour 3548 from the BOV3 family	154
Figure 6.1. Three patterns of non-epidermolytic palmoplantar keratoderma	163
Figure 6.2. Mapping of diffuse NEPPK to the Type II keratin gene cluster	166
Figure 6.3. Mapping of Focal NEPPK (without associated Oesophageal Cancer) to the Type I keratin gene cluster	170
Figure 6.4. Mapping of focal NEPPK/oesophageal cancer in Pedigree 2216 to the TOC region	174
Figure 6.5. Summary of recombination events identified at the 17q24-25 TOC locus in the American focal NEPPK/oesophageal cancer pedigree, 2216	180
Figure 6.6. Examples of allele status in the chromosomal region 17q24-25 (TOC locus) in oesophageal carcinomas	184
Figure 6.7. 17q24-25 characterisation of somatic cell hybrid DNA	189
Figure 6.8. Somatic cell hybrid, YAC and BAC characterisation with D17S1603 and D17S801	193
Figure 6.9. Envoplakin (EVLP), a candidate for the TOC gene?	196
Figure 6.10. Genetic exclusion of the type I keratins with Punctate palmoplantar keratoderma segregating in pedigree 2039	200
Figure 7.1. Hypothetical model for a hormonally regulated dosage effect of BRCA1 (and BRCA2) in sporadic breast cancer	208
Figure 7.2. Two models implicating EVLP in Oesophageal Cancer susceptibility	213

TABLES

Table 1.1. Factors associated with breast cancer risk	12
Table 3.1. Two-point LOD scores for 17q DNA marker loci segregating in 16 breast cancer families	72
Table 4.1. Summary of marker loci content of YACs in the BRCA2 region by PCR amplification	111

Table 5.1. Haplotype comparison of the three ICRF breast-ovarian cancer families carrying the 185delAG mutation in BRCA1	143
Table 5.2. 17q allele loss in sporadic breast tumours	149
Table 5.3. Similar patterns of allele loss in grade 3 breast tumours from BOV3	153
Table 5.4. Distribution of BRCA1 and BRCA2 allele loss among 45 sporadic grade 3 infiltrating ductal breast carcinomas	156
Table 6.1. Linkage of diffuse NEPPK to the Type II keratin gene cluster at 12q11-13	161
Table 6.2. Linkage of focal NEPPK to the Type I keratin gene cluster at 17q12-21	168
Table 6.3. Linkage analysis of focal NEPPK with microsatellite marker loci mapping in the region harbouring the TOC locus in pedigree 2216	177
Table 6.4. Disease haplotype comparison of the three TOC-linked focal NEPPK/oesophageal cancer pedigrees	181
Table 6.5. 17q allele loss in sporadic oesophageal carcinomas	185
Table 6.6. LOD scores obtained for pedigree 2039: evidence against linkage with punctate PPK and candidate chromosomal regions	198

ABBREVIATIONS

APC	Adenomatous polyposis coli
APS	Ammonium persulphate
AT	Ataxia telangiectasia
BAC	Bacterial Artificial Chromosome
bp	Base pairs
BRCA1	Breast Cancer susceptibility gene locus 1
BRCA2	Breast Cancer susceptibility gene locus 2
BSA	Bovine serum albumin
cDNA	Complementary DNA
Ci	Curies
cpm	Counts per minute
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytosine triphosphate
dGTP	Deoxyguanosine triphosphate
dTTP	Deoxythymidine triphosphate
dNTP	dATP, dCTP, dGTP or dTTP
ddATP	Dideoxyadenosine triphosphate
ddCTP	Dideoxycytosine triphosphate
ddGTP	Dideoxyguanosine triphosphate
ddTTP	Dideoxythymidine triphosphate
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
EVLP	Envoplakin
g	Grammes
g	Gravitational force
HGMP	Human Genome Mapping Centre
IPTG	Isopropyl- β -D-thio-galactopyranoside
kb	kilobase pairs
KRT	Keratin
L	Litre
LOD	Log ₁₀ of the odds ratio
LOH	loss of heterozygosity
Mb	Megabase

M	Molar
µg	Microgramme(s)
mg	Milligramme(s)
min	Minute(s)
µl	Microlitres(s)
ml	Millilitre(s)
mM	Millimolar
MMTV	Mouse mammary tumour virus
NEPPK	Non-Epidermolytic Palmoplantar Keratoderma
ng	Nanogramme(s)
°C	Degrees Celcius
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PPK	Palmoplantar keratoderma
Rb	Retinoblastoma gene locus
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
s	second(s)
SDS	sodium dodecyl sulphate
TAE	Tris, acetate, EDTA
TBE	Tris, borate, EDTA
Temed	NNN'N'-tetramethylethylene-diamine
TOC	Tylosis and Oesophageal Cancer locus
Tris	Tris (hydroxymethyl) amino-methane
v	Volts
VNTR	Variable number tandem repeat
X-gal	X-galactosidase (5-bromo-4-chloroindoyl-β-galactosidase)
YAC	Yeast Artificial Chromosome

CHAPTER 1: INTRODUCTION

This thesis describes the genetic analysis of families in which a susceptibility to certain forms of malignancy appears to be segregating. Initial studies were directed towards the genetic analysis of breast and breast-ovarian cancer families.

1.1.Breast Cancer

The most frequently occurring malignancy in women, excluding skin cancer, is breast cancer (CRC Cancer Surveys). In the United Kingdom, 25,000-30,000 new cases are diagnosed each year. To date, the most effective treatment of the disease is surgery followed by local radiotherapy or whole body chemotherapy. Survival is greatly diminished if the cancer has metastasised. Therefore, early detection of the disease is a priority for improved survival rates.

1.2. Risk Factors

A number of factors have been shown to be associated with increased risk of developing breast cancer (Table 1.1), some of which are discussed.

Table 1.1. Factors associated with breast cancer risk (adapted from Petrakis *et al.*, 1982)

Increased risk	Decreased risk
over age 40 years	racial origin, Asian
high socio-economic status	oophorectomy before age 40
early age of menarche	pregnancy before age 21
family history	
first pregnancy over age 30	
early exposure to x-irradiation	
high fat diet	

1.2(i) age

The risk of developing the disease increases dramatically with age; the majority of breast cancer cases occur in women around 70 years of age (CRC Cancer Surveys). This is probably related to longtime exposure to environmental/hormonal risk factors and the time required for a breast cell to progress to the cancerous state.

1.2(ii) reproductive behaviour

Breast cancer risk appears to be associated with prolonged exposure to the hormone oestrogen. For example, women who have an early menarche and late menopause are at greater risk of breast cancer (Kampert *et al*, 1988). Also, oophorectomy has a protective effect as does an early age of full term pregnancy. In the normal breast, oestrogen induces the epithelial stem cells to divide but only in combination with other endocrine products such as insulin, thyroxine, progesterone, prolactin and growth hormone. Therefore, a correct balance of these complex hormonal interactions is required for normal breast epithelial growth.

Due to the association of oestrogen with breast cancer risk, clinical trials are in progress with the anti- oestrogen, tamoxifen, in an attempt to counteract the growth stimulatory activity of the hormone. However, altering the levels of oestrogen can perturb the balance of hormones that it interacts with and may have a disadvantageous effect in some women. For example, there is a small increased risk of endometrial and uterine cancer associated with tamoxifen treatment (Seoud *et al*, 1993).

1.2(iii) obesity

The risk of developing breast cancer has been associated with a high fat intake/ or high body mass, in particular, abdominal obesity in postmenopausal women (reviewed in Stoll,1994). This may be related to high levels of oestrogen exposure derived from the aromatisation of androgen from the fat deposits (Kampert *et al*, 1988). No increased risk has been found in obese premenopausal women (Albanes, 1987). The lower

breast cancer incidence observed in vegetarian women would support a role for obesity in breast cancer risk as, generally, this population are slightly less obese and have a reduced fat intake than omnivorous women (Phillips, 1975).

Predisposition to abdominal fat deposition may be genetically linked to breast cancer susceptibility (Sellers *et al*, 1992; 1993). Women usually deposit fat on their hips whereas women who deposit fat in the abdominal region may be genetically predisposed to do so, though a fat-rich diet clearly is an important factor (Sims, 1990). Alternatively, the hormonal changes that occur with abdominal obesity may elevate breast cancer development in genetically susceptible women.

1.2(iv) country of birth/residence

The continent which shows the lowest incidence of breast cancer is Asia (Petrakis *et al*, 1982; Aaltomaa *et al*, 1992). This, in part, may be related to the population having a lower fat diet and poorer socio-economic climate than Western Europe and the US. Environmental factors clearly have a role as demonstrated by migrant studies. Women born in low breast cancer risk countries, but brought up in a high risk one, show an increased rate of breast cancer incidence. For example, the female Japanese population, who emigrated to Hawaii 2-3 generations ago, have a similar risk of developing breast cancer as that of the indigenous female Hawaiian population (Aaltomaa *et al*, 1992).

1.2(v) family history

The clearest risk factor for breast cancer development is having a family history of the disease. Although, approximately 95 % of breast cancer cases are thought to be sporadic, that is, they occur without a highly penetrant genetic component, a small proportion of cases arise from the inheritance of a mutation in genes which confer an extremely high risk of disease development. There have been numerous epidemiological studies in which a number of breast cancer cases have been shown to occur in families. This can be explained either as chance clustering of sporadic disease

or due to an inherited susceptibility (Bishop *et al*, 1988; Claus *et al*, 1991; Slattery and Kerber, 1993). In a proportion of these families, there is clear evidence of a dominant autosomally inherited susceptibility gene segregating through the generations with the highest proportion of genetic cases occurring at a young age, often with bilateral disease.

Numerous studies of familial risks of breast cancer have been conducted to estimate the risks to female relatives of affected individuals, particularly those diagnosed at an early age, of developing the disease (for example, Slattery and Kerber, 1993). The most extensive study was the Cancer and Steroid study (CASH) which was performed on 4730 cases of breast cancer diagnosed between the ages 20 and 54 by comparing family history to a large population of matched controls (Claus *et al*, 1990). This, and other studies demonstrated a familial affect on breast cancer risk to relatives of affected individuals. From the CASH study, the age of onset of breast cancer in patients dramatically affected the relative risk of female family members developing the disease. The relative risk to first degree female relatives was increased less than twofold if the disease was diagnosed over age 50, but increased to fivefold with cases diagnosed below age 40. The risk of breast cancer in women also increased if they had two or more first or second degree affected relatives. A further indication of a genetic component in a proportion of breast cancer cases was demonstrated by the observation that relatives of individuals with bilateral disease were at greater risk than relatives of unilateral cases (Bernstein *et al*, 1992).

Epidemiological studies in relatives of breast cancer cases which are likely to have a genetic component have demonstrated increased risks of the development of other cancers. One genetic association is that of breast cancer with ovarian cancer. A moderate increase in risk of ovarian cancer in relatives of women with breast cancer and *vice versa* has been observed (Schidkraut *et al*, 1989). A number of families in which breast and ovarian cancer segregates through consecutive generations have been reported suggesting the existence of a gene(s) which confers susceptibility to both

diseases (Lynch *et al*, 1978). A second familial association is that of early onset breast cancer with childhood bone and soft tissue sarcoma, termed the Li Fraumeni cancer syndrome (reviewed Birch, 1990). Another cancer syndrome in which breast cancer is a component is Cowden (or multiple hamartoma) disease in which thyroid cancer, digestive tract carcinomas and breast cancer may occur at a relatively early age (reviewed Hanssen and Fryns, 1995). Other reports, derived either from population studies or extensive cancer families, have implied possible genetic associations of breast cancer with prostate cancer (Cannon *et al*, 1982; Tulinius *et al*, 1992; Arason *et al*, 1993) and colorectal cancer (Phipps and Perry, 1989).

Breast cancer risk is increased in relatives of individuals affected with the autosomal recessive progressive neurological disorder, ataxia telangiectasia (AT) (Swift *et al*, 1992). These presumed AT heterozygotes are also at increased risk of developing other cancers, for example, stomach cancer and leukemia. Estimates suggest the contribution of AT heterozygotes to breast cancer incidence could vary from a minimal to a significant proportion (Easton *et al*, 1993a). The recent cloning of the AT gene, ATM, will allow the molecular identification of AT heterozygotes and a greater understanding of its role in breast cancer predisposition (Savitsky *et al*, 1995).

1.3. Segregation Analysis

Segregation analyses have tested a number of genetic models to support the existence of inherited predisposition for breast cancer. Using this analysis, it is possible to fit the best single locus genetic model to the susceptibility gene by varying parameters such as population allele frequency, penetrance (the probability that an individual with or without the gene defect will develop the disease by a given age) and mode of transmission. Using the CASH data set, the most likely model to explain familial breast cancer is that susceptibility is conferred by an autosomal dominant gene with a population frequency of 0.0033 (Claus *et al*, 1991). For disease gene carriers, the cumulative risk or penetrance of breast cancer was estimated to be 38% by age 50 and 67% by age 70. For non-gene carriers, the cumulative risk was estimated to be 1.5%

by age 50 and 5% by age 70. Using this model, about 35 % of breast cancer cases diagnosed below age 30 were estimated to be due to a defect in this gene, compared with about 1% of cases diagnosed over age 80 (Claus *et al*, 1991). Similar models have been found for other data (Bishop *et al*, 1988; Newman *et al*, 1988; Iselius *et al*, 1991). This gene(s) is unlikely to account for the majority of breast cancer cases which will occur sporadically, although it is also possible that there are lower penetrance genes in the population which “modify” the risk of breast cancer. These “modifying” genes may have a role in a much larger number of cases.

1.4. Linkage Analysis

Segregation studies support the presence of one (or more) major susceptibility gene(s) in familial breast cancer. Using the families which have a number of affected cases and show evidence of autosomal dominant inheritance, genetic linkage studies can be performed. These studies aid in the identification of disease gene(s) by pinpointing the position in the genome to which it is localised. Linkage analysis is an extremely powerful tool and has been used successfully to map numerous disease genes of either dominant or recessive inheritance with complete or variable penetrance. This strategy has enabled the identification of genes for heritable traits and diseases for which only the phenotype was known and not the protein product (reviewed Ballabio, 1993).

1.4.1. Concept of Genetic Linkage

The concept of genetic linkage is based on various observations regarding features of heredity and the extensive genetic variability found in humans (for a review see Strachan and Read, 1996). Key observations include:

- (i) two traits are more likely to be coinherited if their corresponding genes are close together on the same chromosome, termed non- independent segregation.
- (ii) if two traits map to different chromosomes, they segregate independently of each other because of the independent segregation of the chromosomes in meiosis
- (iii) the more distant two traits map from each other, the more likely that they show independence of segregation. This is due to the process of genetic exchange which

occurs at meiosis between paired chromosome homologues. The greater the distance between two loci, the greater the likelihood that recombination occurs between them. The position and frequency of these cross-overs differ from one meiosis to the next, with on average one or two exchanges per chromosome per meiosis.

1.4.2. Basic mathematics of linkage analysis (see Ott, 1992)

The basic requirements for a linkage study are an understanding of the mode of inheritance of the trait or disease under investigation in families and the availability of genetic markers (see section 1.5). Additionally, the sample size has to be large enough that the statistical evidence of linkage can be found. Problems encountered with certain modes of inheritance for specific traits are discussed in the next section. The development of genetic markers and their integration into a genetic linkage map is discussed in section 1.5.

Even for two adjacent loci, genetic linkage is not complete as there is always a possibility of recombination between the two 'linked' alleles at each locus during meiosis, though the frequency is much less than that between two distant loci. The probability that a cross-over occurs between two loci is referred to as the recombination fraction, represented by theta (θ), and is calculated from the proportion of recombinants generated as a fraction of the combined number of meioses. The recombination fraction is thus an indication of genetic distance between two loci. For two unlinked loci, that is, when alleles of the two loci segregate independently of each other, the value for θ is close to 0.5 and would indicate the two loci are far apart on the chromosome or on different chromosomes. For two loci mapping very close to each other on the chromosome, the theta value would be nearer to zero as there would be fewer recombination events observed between alleles at the two loci. Genetic distance between loci is measured in centiMorgans (cM): a Morgan is the genetic distance in which exactly one recombination event occurs. On average over the whole genome, one cM approximates to the physical length of one Mb. However, there is not an exact relationship between the two measures of distance as there are differences in

recombination frequencies between chromosomal regions and the sexes for specific chromosomes.

In human linkage studies, the LOD score method is used to investigate linkage (Morton, 1955). It measures the relative likelihood of observing a given distribution of a trait/disease and a genetic marker segregating within a family for a range of recombination fractions. The relative odds of the trait/disease locus and the marker locus being linked against no linkage are expressed as the \log_{10} of the odds ratio or the 'LOD score' for each possible recombination fraction. The maximum value of the LOD score (over different recombination fractions) can be incorporated into a statistical test for linkage. Satisfactory evidence for two loci being genetically linked is generally accepted as a LOD score of 3.0 or greater. A LOD score of 3.0 at a recombination fraction of 10 % indicates that the segregation of alleles for a trait with a genetic marker is one thousand times more likely to have occurred if the two loci are genetically linked (that is, recombination between the two loci occurred only once in ten meioses) than if they were genetically distant. For any two randomly chosen loci, the odds for those two loci being linked is estimated to be 1: 50 (Elston and Lange, 1975). The posterior odds for linkage between two loci when a LOD score of 3 is attained is calculated from multiplying the odds from the data, 1000: 1, with the prior odds of linkage, 1: 50 which is 20: 1. This equates to $p = 0.05$, the conventional threshold for statistical significance. Thus the posterior probability of the two loci being linked is 95 %, and that of a false positive result being 5 %. The greater the LOD score, the lower the probability of a false positive, for example, a LOD score of 5 will be misleading, on average, one time in 2,000.

A LOD score of - 2 is a strong indicator of exclusion of linkage between two loci (Morton, 1955). Those scores between -2 and +3 are generally regarded as inconclusive. Though in some circumstances, lower LOD scores than 3 could be suggestive of linkage when there is prior evidence of a disease locus mapping to the region.

1.4.3. Problems encountered in a linkage study

The LOD score method can be implemented to cope with the complexities of mapping traits/ diseases including problems encountered with untyped individuals, potential phenocopies and variable penetrance. The majority of these variables can be combined in the analysis by the use of computer algorithms (for example, Lathrop and Lalouel, 1984). Some of the problem areas, particularly those related to breast cancer inheritance, are discussed below :

1.4.3(i) Genetic heterogeneity

As the probability of linkage is measured logarithmically, LOD scores for two or more families with the same trait can be added together to increase the power of detecting linkage of the trait with a genetic marker. Therefore analysis of small families (rather than one extensive pedigree) can be used in linkage mapping. However, genetic heterogeneity of a disorder can create difficulties in the interpretation of data when combining LOD scores, for example, if mutations in two or more distinct gene loci predispose individuals to apparently identical disease phenotypes. Two ways of limiting this problem are :

(a) analyse families from isolated geographic communities for a disease which is sufficiently rare enough that mutations in two distinct gene loci are unlikely to be segregating through the population

and/or

(b) analyse a few extended pedigrees when available. However, if the disease mutations are relatively common in the population, it is possible that distinct mutations may be segregating in different branches of the family.

1.4.3(ii) incomplete penetrance

Even though an individual may have inherited a mutation in a susceptibility locus, other factors may affect the clinical presentation of the disease such as environmental effects, chance and/or modifying gene loci. This is of particular relevance to inherited breast cancer. All three factors may have a role in determining whether a woman

carrying a mutation in a breast cancer susceptibility gene develops the disease early in life or, indeed, at all.

1.4.3(iii) Phenocopies

This occurs when individuals present with the disease though they are not carrying a susceptibility mutation. This “sporadic” disease may be explained by a non-inheritable occurrence or by a distinct mutation in the same or different gene to that inherited by the majority of affected cases in the family. As long as the proportion of phenocopies to gene carriers in a family is low enough, linkage analysis models can cope with this problem. The same problem as sporadic disease can occur because of misdiagnosis. The affected status of each individual must be clinically confirmed.

1.5. Genetic Markers

In order to localise a disease gene by linkage analysis, chromosomal homologues must be 'tagged' as they segregate through pedigrees to see if a particular chromosome region follows the inheritance pattern of the disease. This can be achieved by using the extensive genetic variation found in humans visualised in the form of genetic polymorphic markers. A genetic marker must show variation in the form of alleles and each of these morphs must be relatively easy to detect. The degree of polymorphism is indicated by the number and frequency of alleles at each marker locus in the population. An individual is more likely to be heterozygous at a locus if that locus displays a wide range of alleles of equal frequency. It is important that these alleles are relatively stable in the population as the appearance of 'new' alleles can affect the tracing of the marker through a pedigree. Theoretical examples of how variation at a locus can affect the 'tagging' of chromosomes are demonstrated in Figure 1.1.

Initial studies in human genetics were hampered by the impossibility of controlled matings in humans compared to other organisms, such as *Drosophila* and mice, and the lack of genetic polymorphisms which could make existing human matings for specific traits informative for linkage analysis. Therefore, to fully utilise the family

material available to study human inheritance, a comprehensive linkage map of the human genome was required. Work over the last ten to fifteen years has made the prospect of a complete linkage map of the human genome a reality. For example, a genetic map has been constructed which spans a sex-averaged distance of 3,699 cM (Dib *et al*, 1996). This means that there are sufficient genetic marker loci across every chromosome to potentially identify linkage to any gene of interest. The map has developed to the degree that once linkage to a marker has been identified, the chromosomal location is immediately known. The size of the region may be reduced further by screening with close flanking markers to aid in the physical identification of the gene (discussed in section 1.11). However, this is dependent on the presence of informative recombinants in families.

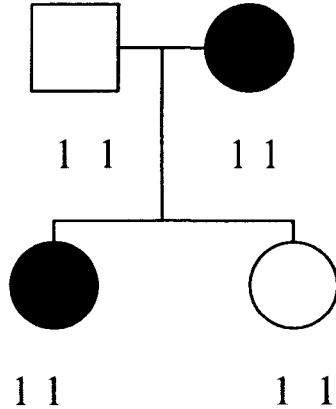
1.5.1. Development of the Human Genetic Linkage Map

Morgan (1910) and colleagues introduced the concept of genetic linkage maps for ordering genes along a chromosome from studies on the fruitfly, *Drosophila melanogaster* (reviewed in Strachan and Read, 1996). One of the first polymorphic systems used in humans were protein polymorphisms, for example, blood group antigens, but there were few available, and they covered only small intervals of the genome (reviewed White and Lalouel, 1988). The breakthrough came in the late 1970's with the detection of variation in DNA sequences. These first DNA marker systems were polymorphisms generated by a single base change or deletion which affected a restriction site. The majority of these marker loci have two alleles; one with the restriction site and the other without. DNA extracted from individuals is enzymatically digested with the specific restriction endonuclease that recognises the restriction site. The digested DNA is resolved by electrophoresis through agarose gels and transferred to a nitrocellulose or nylon membrane (Southern, 1975). A DNA probe spanning the present/ absent restriction site is then hybridised to the genomic DNA and the allelic sequence variation visualised. Such polymorphisms are referred to as restriction fragment length polymorphisms (RFLPs) (Botstein *et al*, 1980). This DNA variation was found to be sufficiently common in the human genome to suggest

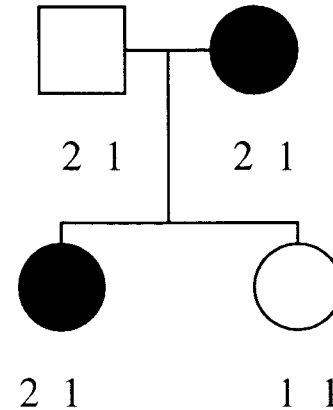
that a complete linkage map was possible (Willard *et al*, 1985). Genetic maps were constructed using these RFLP markers which consisted of linkage groups on all chromosomes (Donis-Keller *et al*, 1987). A set of three generation families collected by the Centre d'Etude du Polymorphisme Humain (CEPH) were used to order these markers by multipoint linkage analysis. This analysis computes LOD scores for the most likely location of the marker locus based on its segregation in families in relation to other marker loci, for example, using the LINKAGE program (Lathrop *et al*, 1986; Lathrop and Lalouel, 1988). DNA samples from the CEPH families are screened by groups wishing to integrate new genetic markers into the map. RFLP markers enabled the chromosomal localisation of a number of disease genes, for example, the Huntington's disease locus (Gusella *et al*, 1983).

The majority of the RFLP DNA markers are relatively uninformative exhibiting a low level of polymorphism (maximum heterozygosity 50%) and, as a consequence, are often uninformative for genetic linkage analysis in most families (Figure 1.1.(i) and (ii)). However, a small number of DNA marker loci were cloned which exhibited a far higher degree of allelic variation with heterozygosity levels approaching 90% for some loci. This high level of polymorphism was found to be due to short DNA sequences (11 to 60 bp) being tandemly repeated; such loci are referred to as “minisatellites” or variable number of tandem repeats (VNTRs: Jeffreys *et al*, 1985). These multiallelic systems are ideally suited to linkage studies as individuals are likely to be informative or heterozygous at the locus (Nakamura *et al*, 1987; Figure 1.1. (iii)). The application of the polymerase chain reaction (PCR: Saiki *et al*, 1988) increased the speed of DNA marker analysis plus reduced the quantity and quality of DNA required. Short oligonucleotide primers are designed in unique DNA sequence either side of the repetitive sequence and are used to amplify the DNA sequence in between using a thermostable Taq polymerase. The PCR amplification products are resolved by gel electrophoresis to detect the length variation. However, many minisatellite alleles are beyond the effective limits for PCR (>10 Kb in length).

(i) two allele marker system : one common allele



(ii) two allele marker system : approximate equal allele frequency



(iii) multi allelic marker system: equal allele frequency

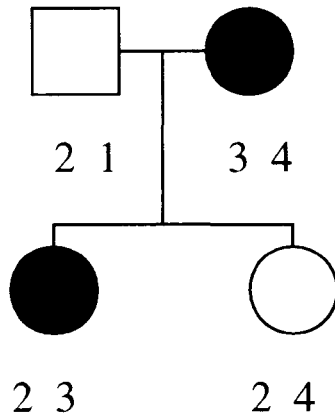


Figure 1.1. The affect of allele number and frequency in linkage analysis

Circle indicates female and square indicates male; if blackened, indicates affected individual. Genotype of each individual for each marker system shown directly under the individual.

(i) As both parents are homozygous for the same allele, it is impossible to determine from which parent, the affected daughter inherited each allele from.

(ii) As both parents are heterozygous for the same two alleles, it is impossible to determine from which parent, the affected daughter inherited each allele from. However, the unaffected daughter clearly has inherited allele 1 from each parent, inferring that the disease locus segregates with allele 2.

(iii) Marker is fully informative as parents are heterozygous for different alleles. Affected daughter has clearly inherited allele 3 from her affected mother, whilst her unaffected sib inherited the maternal allele 4.

A major disadvantage of VNTR markers for linkage analysis is that the majority are clustered towards the telomeres of the chromosomes (Royle *et al*, 1988). Thus a genetic map of the genome could not be based on this marker system alone. Another class of tandemly repeated DNA sequences are found in the genome. These are in great abundance and it is estimated that a repeat block is present every 30-60 kb with a relatively even distribution throughout the genome (Hamada *et al*, 1984). This repetitive DNA family consists of short dinucleotide repeats such as (dC-dA)_n, (dG-dT)_n, which are designated “CA” repeat microsatellites. They are reported to have arisen due to “slippage” in DNA replication and are often highly polymorphic with multiple alleles due to length variation of the “CA” blocks (Hearne *et al*, 1992). Weber and May (1989) demonstrated the usefulness of these microsatellites as an abundant source of genetic markers. Generally, the longer the run of repeats, the greater the number of alleles at that locus (Weber, 1990). The properties that have made this marker system so useful for linkage studies are as follows:

- (i) frequent distribution in the genome
- (ii) high polymorphic multi-allelic systems
- (iii) rapid and easy detection by PCR
- (iv) ability to PCR amplify DNA from archival pathological material, for example, to increase the informativeness of families for linkage analysis.

Other classes of polymorphic markers have been developed, for example, the polydeoxyadenylate tract found in *Alu* repetitive elements (Economou *et al*, 1990), but the majority of marker loci which form the basis of current genetic linkage maps are the microsatellites.

Vast numbers of microsatellite marker loci have been isolated across the genome and integrated into high resolution genetic maps (NIH/CEPH Collaborative Mapping Group, 1992; Weissenbach *et al*, 1992; Gyapay *et al*, 1994, Dib *et al*, 1996). The application of fluorescence-based technology has improved methods of detection of these marker loci (Ziegle *et al*, 1992; Mansfield *et al*, 1994). Sets of microsatellite

marker loci tagged with either one of three fluorescent dyes and with different allele size ranges are available which span each chromosome (for example, Reed *et al*, 1994). The resulting PCR products are pooled and run on the same gel to increase the speed and accuracy of a genomic search in a disease gene linkage mapping study.

1.5.2. Integrated Genetic and Physical map of the human genome

The generation of a comprehensive genetic linkage map of the genome will increase the potential power of disease gene localisation by linkage analysis. However, the actual identification of the gene requires the physical positioning of transcribed sequences in relation to the flanking genetic markers. Such candidate transcribed sequences can then be characterised and mutation analysis performed to identify the disease gene. This approach has been termed “positional cloning” (reviewed Ballabio, 1993). One of the first successes using this approach was the identification of the Cystic Fibrosis gene (Riordan *et al*, 1989).

Numerous methods of identifying transcribed sequences have been used including CpG island mapping (Cross and Bird, 1995), exon trapping (Buckler *et al*, 1991) and the sequencing of randomly selected cDNAs (Adams *et al*, 1991; 1995). Transcribed sequences can be rapidly assigned to chromosomes and physically ordered using, for example, human rodent somatic cell hybrid mapping panels (reviewed Spurr *et al*, 1996), in situ hybridisation (FISH: Buckle and Craig, 1986) and 'fingerprint' contig construction using yeast artificial chromosomes (YACs) (Bellanne-Chantelot *et al*, 1991; Cohen *et al*, 1993). A YAC is a vector in which large fragments of human DNA (up to 1000 kb or greater) can be cloned (Burke *et al*, 1987; Anand *et al*, 1989). The development of YAC contigs has proved the most successful method towards the integration of the physical and genetic maps (reviewed Schlessinger, 1990). Large regions of every chromosome have been covered in YAC contigs facilitating the ordering of any transcribed sequence or any DNA landmark in relation to genetic marker loci (Chumakov *et al*, 1995).

CANCER GENETICS

1.6. Cancer is a multi-step genetic disease

There is strong evidence that a cancer arises from a single cell as a result of clonal expansion (Nowell, 1976; Ponder and Wilkinson, 1986). A focus for molecular biology has been the identification of the genetic changes occurring in that cell which disrupt the controlling mechanisms that prevent every cell becoming cancerous. The model that fits best is that of multistep carcinogenesis (Foulds, 1957; Nowell, 1976). This is explained by the acquisition of changes in genes which control proliferation and cell-cell interactions. Accumulation of these specific gene mutations will allow the “cancer cell” to progress along the tumorigenic pathway enabling it to “escape” the confines of normal cell regulation and metastasise. The hypothesis that these genetic changes are ordered in the multi-step process is clearly demonstrated in the colon cancer pathway (Fearon and Vogelstein, 1990) shown in Figure 1.2; and is supported by the following observations:

1.6(i) Histopathological

From looking at specific tumours, there is evidence for clear histopathological differences in the development of tumours (Foulds, 1957). One of the best examples is the progression of a normal colon epithelial cell to that of a metastasising carcinoma; whereby the cell passes through a hyperproliferative stage, via various classes of adenoma to that of a carcinoma (Sugarbaker *et al*, 1985; Vogelstein and Kinzler, 1993). This pathway is shown in Figure 1.2.

1.6(ii) Epidemiology

In cancer syndromes, every cell has inherited a mutation in the germline, but not all these cells become cancerous. Therefore, other mutational events are required (Hansen and Cavenee, 1987). Each “growth-advantageous” mutation will allow that cell to divide more freely thus increasing the relative proportion of cells carrying that mutation. This increases the probability that one of these mutated cells will acquire subsequent growth selective genetic alterations.

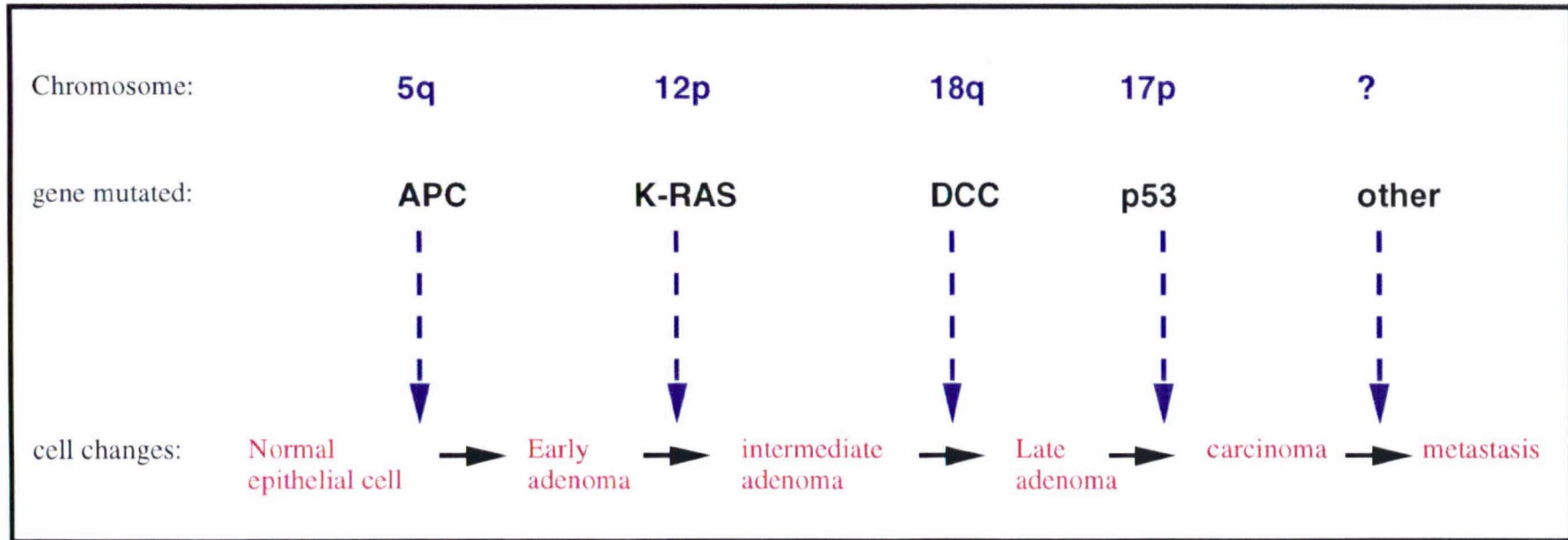


Figure 1.2. Multistep model for colorectal tumourigenesis (adapted from Fearon and Vogelstein, 1990)

Model showing some of the identified genetic changes occurring during the progression of a colon epithelial cell to a metastising cancer cell.

1.6(iii) Non-random genetic aberrations

Cancer cells have abnormal chromosomes. These karyotypic changes are often unevenly distributed throughout the genome (Mitelman and Levan, 1981; Cavenee and White, 1995). Some are common in particular tumour types, for example the Philadelphia chromosome in chronic myeloid leukaemia (CML), a chromosome 9 and 22 translocation $t(9;22) (q34;q11)$ (Prakash and Yunis, 1984). These are termed primary chromosome aberrations. Others are secondary karyotypic aberrations which occur frequently in different tumours of the same cellular origin and are implicated in the progression of that tumour type. The most recurrent karyotypic abnormality observed in breast tumours involves chromosome 1, though other specific chromosomal abnormalities are also observed (Rodgers *et al*, 1984).

The key somatic genetic changes that occur in the multistep process are as a result of mutations (reviewed Loeb, 1991), possibly as a consequence of an inherited defect, for example, defective mismatch repair genes resulting in an elevated mutation rate (Bodmer *et al*, 1994; Lindahl, 1994; MacPhee, 1995) and/or environmental damage.

1.7. Cancer Genes

A number of genes have been implicated in carcinogenesis. They have been identified as having an important inherited component and/or are somatically altered during tumour formation. These genes fall into broadly two classes which have contrasting mechanisms: tumour suppressor genes and oncogenes. Both have roles in the regulation of cellular growth, differentiation and/ or cell-cell interactions.

1.8. Tumour Suppressor Genes

The loss of specific chromosomal regions is a common event in the majority of tumours (Cavenee and White, 1995). These deletions may result in loss of function of a gene or genes which are involved in the etiology of the cancer. These genes are called tumour suppressor genes, and inactivation of the gene removes its inhibitory role in the regulation of cell growth and differentiation (Friend *et al*, 1986). Hybrid

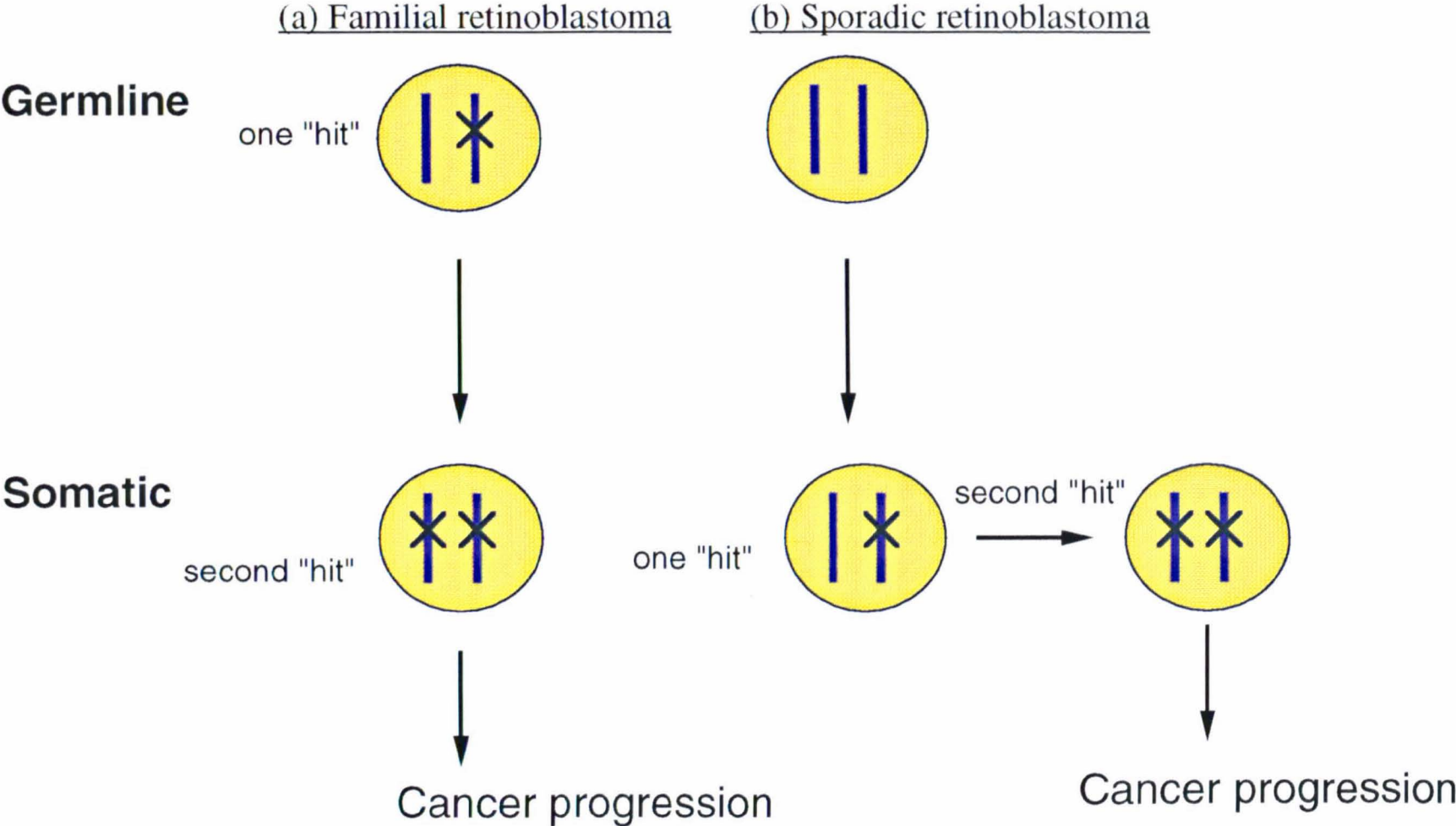
fusion of tumorigenic and non-tumorigenic cells demonstrated that recessive genetic changes were required for tumour formation. These hybrid cells did not give rise to a malignant phenotype unless specific chromosomes were deleted from the non-tumorigenic parent cells (Weissmann and Stanbridge, 1983).

1.8.1. Paradigm of Retinoblastoma

Epidemiological studies on families with the rare childhood cancer, retinoblastoma, led Knudson (1971) to propose the “two-hit” hypothesis for the action of tumour suppressor genes in inherited cancer. Retinoblastomas are malignant tumours that develop in the retina. From the observation that bilateral cases were often diagnosed at an earlier age and had a family history, Knudson suggested that the cancer resulted from two mutational events; one in each allele of the same gene. These bilateral cases were likely to be carrying a germline mutation in the susceptibility gene but required a second somatic mutational event in the normal allele of a retinal cell before the cancer manifested. This hypothesis could also explain sporadic retinoblastoma which was more likely to occur in unilateral cases and at a later age. The same gene was likely to be involved but two somatic mutation events in that gene in the same cell were required. As two hits are required in the tumour suppressor gene, the inherited susceptibility allele is actually recessive at the cellular level. An example of this mechanism is shown in Figure 1.3. The apparent dominant inheritance of cancer can still be explained if the susceptibility allele is recessive at the cellular level as a second hit is highly likely to occur in at least one retinal cell carrying a germline mutation. Studies leading up to and after the cloning of the retinoblastoma susceptibility (RB) gene on chromosome 13q14 supported this mechanism. These include:

- (i) interstitial deletion at 13q14 in germline cells of retinoblastoma cases (Knudson, 1976)
- (ii) cytogenetic deletion of 13q14 in a proportion of retinoblastoma tumours (Benedict *et al*, 1983)
- (iii) linkage analysis of the susceptibility gene in retinoblastoma families to this region (Sparkes *et al*, 1983)

Figure 1.3. Knudson's two hit model for retinoblastoma



(iv) detection of mutations in the retained allele of the RB gene in familial and sporadic tumours (Friend *et al*, 1986).

Using the model of inactivation of tumour suppressor genes, it is possible to identify candidate regions which may harbour these genes in specific familial and sporadic cancers, for example, to target linkage studies. The identification of specific chromosomal abnormalities associated with familial and/or sporadic cancer may be detected cytogenetically and/ or by allele loss studies.

1.8.2. Allele loss

For inherited cancer, the first mutation in the tumour suppressor gene is passed to all cells through the germline. This mutation is often a small DNA sequence change which disrupts the function of the protein translated from that allele, such as an amino acid change causing premature truncation of the protein. Following Knudson's hypothesis a second somatic mutation is required to 'knock-out' the normal allele. This second event is often a deletion of the whole gene, either through loss of the whole chromosome or due to a smaller interstitial deletion. Using polymorphic genetic markers, it is possible to detect this deletion in the tumour (Figure 1.4). If the genetic marker maps within the deleted region and is heterozygous in normal cells from the individual, then only one allele will be present in the tumour. This is termed "loss of heterozygosity" (LOH) or allele loss (Lasko *et al*, 1991). This loss of an allele, if seen frequently in a number of tumours, may indicate the presence of a tumour suppressor gene in the common region of deletion. This method of analysis has been successfully reported for a number of tumour suppressor genes associated with familial cancer, for example, Retinoblastoma, RB1 (Cavenee *et al*, 1983), TP53 (Li-Fraumeni syndrome: Hollstein *et al*, 1991; Donehower and Bradley, 1993), von Hippel-Lindau (VHL) disease (Latif *et al*, 1993) and adenomatous polyposis coli (APC) (Sasaki *et al*, 1989). Using the high density genetic maps available, it is possible to produce LOH patterns or allelotypes for all familial and sporadic tumours (Vogelstein *et al*, 1989). These allelotypes reveal numerous regions of the genome which may harbour tumour

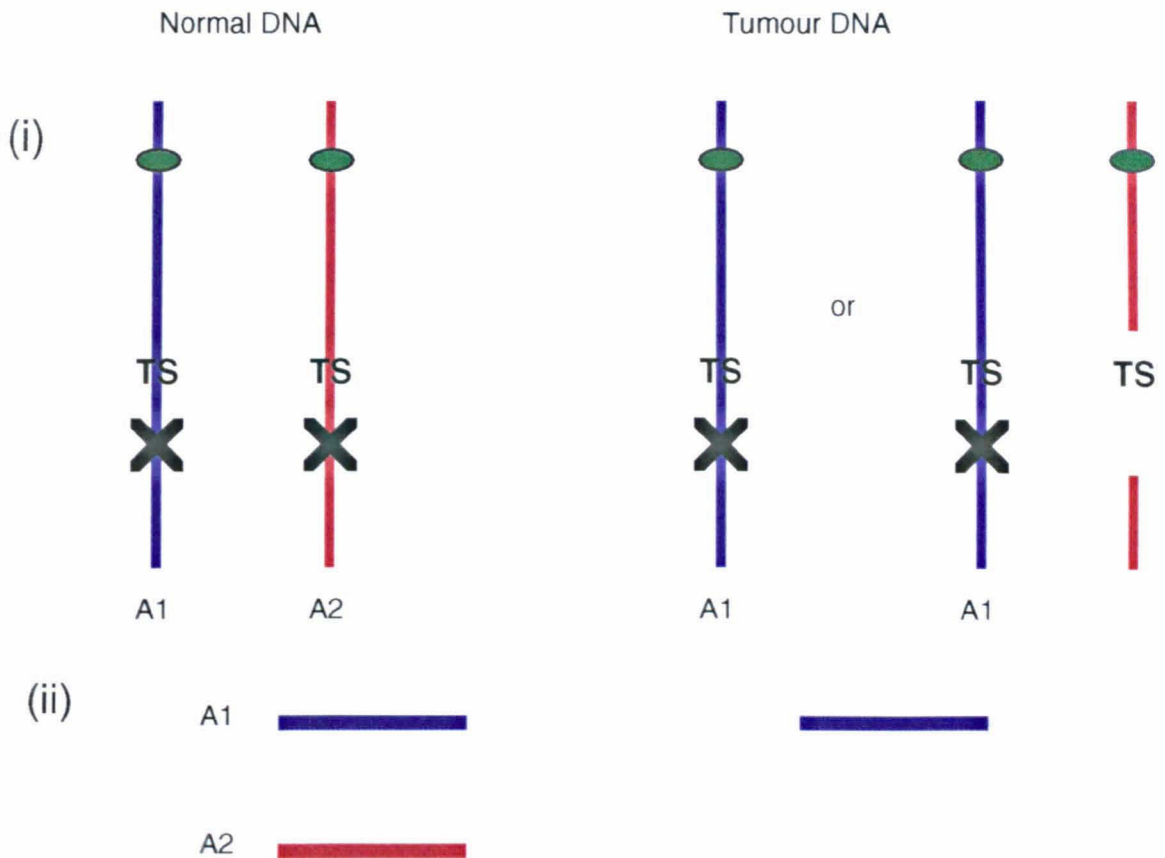


Figure 1.4. Allele loss to detect tumour suppressor genes

This figure demonstrates the detection of allele loss towards the identification of a tumour suppressor gene.

(i) Ideogram showing each chromosome homologue tagged by colour: blue indicates carries allele 1 (A1), red indicates allele 2 (A2). Allele numbering shown under each homologue. The position of the putative tumour suppressor gene (TS).

(ii) Visualisation of the two different sized allele fragments, A1 and A2. Loss of allele 2 (red band) in the tumour cell indicates loss of whole chromosome or an interstitial deletion. Both events delete one copy of the TS gene.

suppressor genes. These deletion maps have also aided in the identification of tumour suppressor genes which apparently are not involved in familial cancer but have a role in later stages of the cancer pathway, for example, the Deleted in Colon Carcinoma (DCC) gene in colon cancer (Fearon *et al*, 1990).

The majority of familial tumour suppressor genes appear to be inactivated at the first step in the cancer pathway. One possible exception may be in familial adenomatous polyposis coli, for which the second inactivating mutation at the APC gene locus may occur after polyp formation (Solomon *et al*, 1987). This means the inherited mutant APC allele has a slight phenotypic effect on some of the colonic epithelial cells and therefore, is not completely recessive. The APC protein appears to have multi-functional domains (Grodin *et al*, 1991). One central domain associates with catenins which, in conjunction with E-cadherin, among other proteins, form an important component in cell-cell adhesion and communication (Rubinfeld *et al*, 1993; Su *et al*, 1993). This may explain the proposed dosage effect with polyp formation. A slight reduction in the APC/catenin contact inhibition signal may induce colonic epithelial cells to grow and result in abnormally organised cells or polyps.

1.8.3. Genomic Imprinting

Another mechanism for inactivating the normal inherited allele is by genomic imprinting whereby only one of the parental alleles is expressed, the other silenced by allele-specific CpG methylation (reviewed Squire and Weksberg, 1996). Imprinting has been demonstrated at the Wilms tumour locus in a tissue specific manner (Zhang *et al*, 1993; Jinno *et al*, 1994). Hypermethylation for ‘silencing’ gene transcription has been demonstrated as a mechanism for a number of other putative tumour suppressor genes, including the von Hippel-Lindau (VHL) gene (Herman *et al*, 1994) and the HIC-1 gene (Wales *et al*, 1995). In an analogous fashion to “allelotyping” tumours (see previous section) to identify regions potentially harbouring tumour suppressor genes, a number of chromosomal regions have been demonstrated to be

hypermethylated in specific tumour types, for example, chromosome 18q21 in lung carcinomas (Nagatake *et al*, 1996).

1.8.4. Tumour suppressor genes and breast cancer

From allelotyping and linkage studies, very few tumour suppressor genes have been identified which have a role in breast cancer. Numerous LOH studies have implicated a number of genomic regions which may harbour tumour suppressor genes in familial and sporadic breast cancer (for example, Devilee *et al*, 1989; Lindblom *et al*, 1993). These regions demonstrate allele loss in greater than 20 % of breast tumours and often include multiple sites on various chromosomes. For example, two distinct regions on chromosome 3p (Chen *et al*, 1994) and at least two regions on 17q (Cornelis *et al*, 1993). Genes that have been identified which fulfill the role of tumour suppressors in breast cancer progression include:

1.8.4.(i) E-cadherin

The E-cadherin gene maps to chromosome 16 in the smallest deletion unit identified in breast tumours from allele loss studies, at 16q22.1 (Cleton-Jansen *et al*, 1994). Reduced expression of E-cadherin has been associated with invasive potential (Oka *et al*, 1993). In four lobular carcinomas displaying allele loss of this region, somatic mutations generating premature stop codons were detected fulfilling its putative function as a tumour suppressor (Berx *et al*, 1995). E-cadherin is a member of the cadherin family which are involved in the formation of intercellular junctions (Birchmeier *et al*, 1995). A phenotype of lobular breast cancers is that of scattered tumour cell growth which would correlate with loss of function of a cell-cell adhesion molecule, such as E-cadherin.

1.8.4.(ii) Retinoblastoma (RB)

Inherited mutations in the RB1 gene predispose individuals to predominantly retinoblastoma but are also associated with other malignancies such as osteosarcoma

and to a lesser extent, breast cancer (Friend *et al*, 1986). In sporadic breast tumors exhibiting allele loss at this locus (13q14), very few somatic mutations have been identified in the retained allele (Lee *et al*, 1988; T'Ang *et al*, 1988). This could suggest that another tumour suppressor gene(s) maps to this region which is the target for deletion.

The RB1 protein is a transcription factor and in its unphosphorylated form suppresses S phase and cell growth (Goodrich *et al*, 1991). This protein is expressed in most cell types (Lee *et al*, 1988), but, when mutated in the germline, only a few tumour types develop, mainly retinoblastomas. This would indicate that retinoblasts are highly dependent on RB1 in cell cycle control. For other cell types such as breast epithelial cells, cell cycle control may be less dependant on RB1; RB1 loss may only be selective if mutational events have occurred in other associated cell cycle regulatory genes.

1.8.4.(iii) p53

Germline mutations in the p53 gene have been identified in families with the autosomal dominant Li Fraumeni cancer syndrome in which breast cancer is a component (Srivastava *et al*, 1990; Malkin and Friend, 1993). From these studies, it is estimated that approximately 1% of familial breast cancer cases may occur as a result of inactivation of this gene. High allele loss rates at the p53 locus have been detected in a number of tumour types, including breast tumours (Mackay *et al*, 1988; Devilee *et al*, 1990; Prosser *et al*, 1990). Mutations in the remaining p53 allele of the tumour have been demonstrated supporting a role for this gene as a commonly inactivated tumour suppressor gene in breast carcinogenesis (Osborne *et al*, 1991; Varley *et al*, 1991; Coles *et al*, 1992). Additionally, homozygous null p53 mice develop normally, but have a high rate of spontaneous tumour growth, particularly lymphomas, at an early age (Donehower *et al*, 1992; Jacks *et al*, 1994).

Like RB1, TP53 protein acts as a transcription factor in the G1/S checkpoint in the cell cycle in most cell types (reviewed in Cox and Lane, 1995). This transcriptional

regulatory role involves targetting numerous genes including MDM2, a putative oncogene (Oliner *et al*, 1992), and the cyclin dependant kinase (CDK) inhibitor, CIP1/WAF1/p21 (El-Deiry *et al*, 1993; Xiong *et al*, 1993). One key role of p53 may be in the blocking of replication or inducing apoptosis in a damaged cell after exposure to DNA damaging agents (Kastan *et al*, 1991; Kuerbitz *et al*, 1992, Livingstone *et al*, 1992; Cox and Lane, 1995). Additionally, p53 is implicated in a checkpoint for spindle formation so loss of TP53 function may affect cell ploidy (Cross *et al*, 1995). As a consequence of this multi-functional role of p53, a cell not expressing TP53 may rapidly accumulate the genetic mutations required for it to become malignant.

1.9. Oncogenes

The second major class of genes involved in the carcinogenic pathway are the oncogenes. They differ functionally from the tumour suppressor genes in that a single aberrant allele is dominant at the cellular level, resulting in an overproduced or overactive form of the protein. The normal form of the gene is termed the proto-oncogene, but after mutation, gene fusion (for example, chromosomal translocation: reviewed Sorensen and Triche, 1996) or DNA amplification, an oncogenic form of the gene is produced. Oncogenes stimulate the cell to replicate via a number of growth promoting mechanisms (Cooper, 1990; Cantley *et al*, 1991; Cavenee and White, 1995). They are implicated in the various growth stimulatory pathways of the cell and include: cell surface receptors which bind growth factors, intracellular signal transduction and nuclear proteins involved in cell cycle regulation.

There are two oncogenes which are transmitted through the germline and known to be involved in cancer predisposition:

(i) RET, a tyrosine kinase receptor, which is mutated in the familial cancer syndrome, multiple endocrine neoplasia types, 2A and 2B (Mulligan *et al*, 1993; Hofstra *et al*, 1994).

(ii) CDK4, a cyclin dependant kinase, mutated in familial melanoma (Zuo *et al*, 1995). This gene forms a complex with cyclin D which regulates RB phosphorylation

at the G1 checkpoint in the cell cycle under the mediation of p16^{ink4a} (Serrano *et al*, 1993). Interestingly the inhibitor of the cyclin D/ CDK4 complex, p16^{ink4a}, is also mutated in a proportion of familial melanoma cases but it functions as a tumour suppressor gene (Hussussian *et al*, 1994; Kamb *et al*, 1994).

1.9.1. Oncogenes and breast cancer

None of the oncogenes which have been identified appear to be implicated in breast cancer predisposition, though a number have been demonstrated as components in the breast tumorigenic pathway. These genes have been identified either by their increased expression and/or DNA amplification in tumour cells. Specific regions of the genome are amplified in some breast tumours (Kallioniemi *et al*, 1994), for example, chromosome 20q13 (Tanner *et al*, 1994). Amplicons may harbour specific oncogenes which have a role in the progression of that tumour. Oncogenes implicated in breast carcinogenesis include:

1.9.1.(i) Epidermal Growth Factor Receptor (EGFR: ErbB-1).

A variant form of this receptor is detected in 78% of infiltrating ductal breast carcinomas (Moscatello *et al*, 1995). This alternatively spliced form of the gene has been demonstrated to be oncogenic from transfection studies into rodent fibroblasts (Velu *et al*, 1987). EGFR also can form heterodimers with another member of the receptor tyrosine kinase family, ErbB-2 (Karunagaran *et al*, 1996). ErbB2 is also frequently amplified in breast tumours (Slamon *et al*, 1987). Overexpression of ErbB2 has been associated with poor prognosis (Slamon *et al*, 1989).

1.9.1.(ii) Cyclin D1.

Elevated expression of cyclin D1 mRNA is found in 45-50 % of breast tumours compared to normal breast epithelial cells (Buckley *et al*, 1993; Bartkova *et al*, 1994). Overexpression of cyclin D1 in the mammary cells of mouse mammary tumour virus-cyclin D1 transgenic mice, results in abnormal cell proliferation and a high incidence of adenocarcinoma (Wang *et al*, 1994). Cyclin D1 is a cell cycle regulator essential for

G1 progression by activating cyclin dependant kinases (CDK) (Matsushime *et al*, 1994). These CDKs phosphorylate the retinoblastoma gene product, pRB, thus inhibiting RB's growth suppressive effect (Sherr, 1994). Additionally, G1 arrest and growth inhibition in breast cancer cells by anti-estrogens has been correlated with a decrease in cyclin D1 mRNA levels (Sutherland *et al*, 1993). Therefore, there is strong evidence that cyclin D1 is a major oncogenic factor in breast carcinogenesis. The observation that cyclin D1 induced neuronal apoptosis (Kranenburg *et al*, 1996) cast doubt on its oncogenic role, but the overexpression of cyclin D1 in this study far exceeded that seen in breast tumour cells. This suggests that there may be a selective level of cyclin D1 overexpression in tumorigenesis.

1.10. Objectives of study

In 1990, Hall *et al* provided the first convincing data for linkage of a DNA marker locus with familial breast cancer. The marker locus was D17S74 which mapped to the chromosomal region 17q21. Initial evidence was strongest in early onset breast cancer families. Linkage of this marker with the disease was confirmed in three large breast-ovarian cancer pedigrees (Narod *et al*, 1991). This indicated the familial breast cancer gene mapping to this region may also predispose gene carriers to ovarian cancer. The disease locus was designated "BRCA1" (Solomon and Ledbetter, 1990).

Chapter 3

This chapter describes the genetic analysis of 17q in ICRF UK breast cancer families. The aims being to ascertain families linked to BRCA1, possibly identify key mapping recombinants and analyse candidate genes. Results chapters 4 and 5 describe analyses performed which were related to developments occurring in the field of breast cancer genetics during the course of this study:

Chapter 4

This chapter describes the characterisation of a YAC contig spanning a second locus for breast cancer susceptibility, BRCA2.

Chapter 5

This chapter describes allele loss studies in familial and sporadic breast tumours at BRCA1 and BRCA2.

Chapter 6

A pedigree has been described in which an epidermolytic form of palmoplantar keratoderma (EPPK) was segregating with susceptibility to breast-ovarian cancer (Blanchet-Bardon *et al*, 1987). This chapter describes the genetic analysis of families with various forms of non-epidermolytic palmoplantar keratoderma (NEPPK). In a proportion of these families, the skin disease was segregating with susceptibility to malignancy, particularly squamous cell carcinoma of the oesophagus. The aim of this study was to determine the genetic basis of NEPPK and to assess the relationship between the various forms of NEPPK and cancer susceptibility.

CHAPTER 2: MATERIALS and METHODS

The majority of media and solutions used in this study were provided by ICRF Media Production, Clare Hall Laboratories. Suppliers from which other reagents were obtained are indicated in the relevant sections in this chapter. The majority of molecular biology procedures are referenced in Sambrook *et al* (1989) and Asubel *et al* (1989). The water used for experimental procedures was obtained from a Millipore Reverse Osmosis System.

2.1. Family Ascertainment

2.1.1. Breast and breast-ovarian cancer families

The majority of families in this study are from the Yorkshire region, with a small number from Southampton and Oxford. They were identified, usually through the diagnosis of a new case in a family, by physicians, oncologists and radiotherapists who then contacted the ICRF Genetic Epidemiology Unit, St. James's University Hospital, Leeds. A pedigree worker (either A. Williams, B. Ward or M. Adams) was assigned to obtain informed consent to approach other family members from the initial case (or proband) and to trace the family history. An important criteria for selection of each family was consistency with an autosomal dominant inherited susceptibility to breast or breast-ovarian cancer (determined by Prof. T. Bishop, ICRF Genetic Epidemiology Unit). Sample availability was also an important consideration for informative linkage analysis. If all the above criteria were met, blood samples were obtained from potentially informative and consenting members of the family. In addition, verification of cancer diagnosis in each family was obtained from death certificates, hospital records and/or pathological material. Further information as to the ascertainment of these families can be found in Spurr *et al*, (1993).

2.1.2. Non-epidermolytic palmoplantar keratoderma (NEPPK) families

Families: 2244, 2242, 2039 and 2239 were ascertained from the London Hospital by Drs. H. Stevens and R. Ratnavel (ICRF Skin Tumour Laboratory, The Royal London Hospital, Whitechapel, London), after examination of a proband from each family. Family 2216 was identified from the literature (Marger and Marger, 1993) and the

pedigree extended with the consent of family members by Dr. H. Stevens. For all families, the NEPPK segregated as an autosomal dominant trait and was 100 % penetrant (though severity was variable between family members: discussed in Chapter 6). As with the families described in section 2.1.1, informed consent was required to contact other affected and unaffected members of the pedigree. Blood samples were taken from all potentially informative family members. For each family member, disease status was carefully determined by either Drs. H. Stevens, I. M. Leigh or R. Ratnavel. A skin biopsy was obtained from the palm from at least two affected members of each pedigree to confirm the non-epidermolytic pattern of the palmoplantar keratoderma by routine electron microscopy (performed by Dr. H. Stevens and colleagues).

2.2. Source of DNA

2.2.1. Breast and breast-ovarian cancer families

Blood samples were taken from potentially informative members of breast and/or ovarian cancer families (ascertained as described in section 2.1.1) by the particular pedigree worker assigned to each family. The 50 ml blood sample was divided into 25 ml of Hepes buffered RPM1 (30 ml of blood) or 4 ml of 0.25M EDTA pH 8.0 (20 ml of blood). This was delivered to the ICRF Human Genetic Resources, Clare Hall. DNA was extracted, for most individuals analysed in this study, from two sources:

(i) T and B lymphocytes were isolated from the 30 ml blood sample on a Lymphoprep gradient (Nycomed) and stored in 10% DMSO in liquid nitrogen. This procedure was performed by B. Laguda (Human Genetic Resources). These lymphocytes were transformed with Epstein Barr virus to produce immortalised B lymphoblastic cell lines (performed by Cell Production, ICRF, Clare Hall Laboratories). For the majority of these cell lines, DNA was extracted using an Applied Biosystems DNA extractor following the manufacturer's recommended protocols by L. Rooke and B. Laguda (both from ICRF Human Genetic Resources). DNA from cell lines of other breast and breast-ovarian family members in the study was extracted using the procedure described in section 2.3.1.

(ii) DNA was also extracted from the 20 ml whole blood sample using the procedure described in section 2.3.2. This was performed by L. Rooke and B. Laguda for the majority of samples.

2.2.2. Non-epidermolytic palmoplantar keratoderma families

For all potentially informative family members, DNA was extracted from 20 ml blood in 4 ml of 0.25M EDTA pH 8.0, as described in section 2.3.2.

2.2.3. Rodent and Somatic Cell Hybrid DNA

DNA samples from mouse, rat, hamster and human/rodent hybrid somatic cell lines for chromosomal assignments (referenced in Spurr *et al*, 1991; Black *et al*, 1993; Kelsell *et al*, 1995a) were extracted by L. Rooke, B. Laguda (both from ICRF Human Genetic Resources) and H. Nicolai (ICRF Somatic Cell Genetics, Lincoln's Inn Fields London).

2.2.4. Paired breast tumour/ lymphocyte DNA

DNA was extracted (by Dr. D. Barnes and colleagues, ICRF Clinical Oncology Unit, Guy's Hospital, London) from 118 paired fresh primary tumour and lymphocyte DNA samples from breast cancer patients attending Guy's Hospital, London. Breast cancer in the majority of these cases was assumed to have arisen "sporadically", that is, without a major genetic susceptibility component.

2.2.5. Archival pathological material

Archival pathological material (excised tumour sections embedded in paraffin wax for long-term storage) was obtained from two sources:

- (i) 10 μ m histological sections of breast and ovarian tumours were obtained from Prof. T. Bishop.
- (ii) 10 μ m histological sections of paired squamous cell carcinomas of the oesophagus (SCCO)/normal oesophageal tissue were obtained from Dr. H. Stevens.

For each tumour sample, a 10 µm histological section was stained with haematoxylin and eosin to confirm the pathology of the tumour (ICRF Histopathology Laboratory, London). This was determined by light microscopy analysis of the stained section under the guidance of Prof. N. Spurr or Dr. H. Stevens. DNA was extracted using the method described in section 2.3.3.

2.2.6. Placental DNA Human placental DNA for competition hybridisation (used in detection of the D17S74 polymorphic system) was extracted by L. Rooke from frozen placentas.

2.2.7. DNA probes

Published polymorphic cloned human DNA probes (D17S74, D17S4, LEW101, LEW102) for southern hybridisation detection of alleles were obtained from the HGMP Probe Resource Centre when based at ICRF Human Genetic Resources. Details of these probes can be found in Easton *et al* (1993b).

2.2.8. Oligonucleotide synthesis

16-25 mer oligonucleotides for polymerase chain reaction (PCR) amplification and for sequencing of DNA were synthesised by the ICRF Oligonucleotide Synthesis Laboratory, Clare Hall Laboratories. Oligonucleotide primers were synthesised using an automated DNA synthesiser (Model 380B, Applied Biosystems), working on cyanoethyl phosphoramidite chemistry. Primer sequences for microsatellite marker loci were obtained from the Genome Database or Genethon, unless otherwise indicated in the text.

2.2.9. YAC and BAC clones

Primary and secondary pools of Yeast Artificial Chromosomes (YACs) and Bacterial Artificial Chromosomes (BACs) containing cloned human genomic DNA inserts were obtained from Research Genetics, Huntsville, USA. Individual YAC clones were obtained from the HGMP Resource Centre, Hinxton Hall, Cambridgeshire.

2.2.10. CEPH DNA

Human DNA samples from lymphoblastoid cell lines derived from three generation families were obtained from the Centre d'Etude du Polymorphisme Humain (CEPH), Paris. Cell lines were maintained and DNA extracted by L. Rooke. First generation members (grandparents) of these families were used for a random control population to estimate allele frequencies at the microsatellite marker locus, D17S1603.

2.2.11 Molecular weight DNA markers

λ DNA digested with *Hind*III (Molecular Weight Marker III) and ϕ x174 DNA digested with *Hae*III (Molecular Weight Marker IX) were obtained from Boehringer Mannheim.

2.3. Extraction of Genomic DNA

2.3.1. Cell lines

Approximately 10^8 cells were washed in PBS (phosphate buffered saline: 140 mM NaCl, 2.6 mM KCl, 10 mM sodium hydrogen phosphate, 1.75 mM potassium hydrogen phosphate) and centrifuged in a 50ml Falcon tube at 1500 rpm for 5 minutes in a Beckman GP Centrifuge. The PBS was removed and the resulting cell pellet was used for DNA extraction.

- 1) The cell pellet was resuspended in 15 ml of Lysis buffer (10 mM Tris-HCl pH 7.6, 100 mM EDTA pH 8.0, 0.5 % SDS: Applied Biosystems) and 100 μ g/ml proteinase K (Boehringer Corporation) before incubation at 50°C for at least 4 hours.
- 2) One volume of a 1:1:1 mixture of phenol:chloroform: H₂O (Applied Biosystems) was added and the mixture centrifuged again at 1500 rpm for five minutes. The top phase was collected. This was performed twice.
- 3) One volume of chloroform was added and the top phase collected after centrifugation again at 1500 rpm for 5 minutes.
- 4) Finally the DNA was precipitated by the addition of 2 volumes of ethanol and 1/10th volume of 2 M sodium acetate. The DNA was spooled out using a plastic inoculating

loop (Nunc). The DNA was dried briefly under vacuum and resuspended at an approximate concentration of 250 µg/ml in H₂O. This solution was then left overnight at room temperature to allow the DNA to dissolve completely.

2.3.2. Whole Blood

DNA was extracted using the Nucleon II kit (for blood volumes of 0.5 ml to 30 ml) from ScotLab using an adapted version of the manufacturer's recommended protocol. A summary of the major steps involved is indicated below (for steps 1-3, cells were kept on ice unless stated):

- 1) Sucrose triton was added to separate red blood cells from lymphocytes. The lymphocytes were pelleted at 1500 rpm for 5 minutes in a Beckman GP Centrifuge.
- 2) Lysis buffer (Applied Biosystems) was added to the pelleted blood lymphocytes. Sodium perchlorate was then added and the cell solution was shaken for 15 minutes at room temperature. The cell solution was then incubated at 65°C.
- 3) Chloroform was added and the cell solution was shaken for 10 minutes at room temperature. After addition of the Nucleon silica suspension, the cell solution was centrifuged at 1500 rpm for 3 minutes.
- 4) The top phase was removed carefully (above the brown Nucleon silica suspension layer) and the DNA was precipitated and resuspended as described in the previous section.

2.3.3. Archival DNA

A method adapted from Smith *et al*, (1992) was used:

Serial 10 µm histological sections were mounted onto glass slides. One slide was stained with haematoxylin and eosin to determine the pathology of the section (as described in section 2.2.5). Tumour (or normal) cells were scraped from slides of sections taken either side of the stained section into a 1.5 ml tube (Eppendorf) using a sterile disposable scalpel. 250 µl of buffer (10 mM Tris-HCl pH 8.0, 0.5 mM KCl, 1.5 mM MgCl₂, 100 µg/ml BSA, 0.45 % Tween and 100 µg/ml of proteinase K) was

added to the tube. Following an overnight incubation at 55°C, the samples were boiled for five minutes.

2.3.4. Yeast DNA (for YACs)

This method was adapted from Asubel *et al* (1989) by Dr. I. Gray (Human Genetic Resources).

- 1) 5 mls of cells in solution were pelleted by centrifugation at 1500 rpm for 5 minutes in a Beckman GP Centrifuge and resuspended in 0.5 ml of sorbitol solution (0.9 M sorbitol, 0.1 M Tris pH 8.0, 0.1 M EDTA) and vortexed.
- 2) 50 µl of 1000 units/ml of lyticase (Sigma) in sorbitol solution and 50 µl of 0.28 M β-mercaptoethanol (Sigma) were added. The solution was vortexed and incubated at 37°C for 60 minutes in a shaking incubator.
- 3) The cells were pelleted by centrifugation again at 1500 rpm for 5 minutes and resuspended in 0.5 ml of 0.5% SDS, 100 mM Tris pH 8.0, 50 mM EDTA. After transfer to a 1.5 ml tube, the solution was incubated at 70°C for 20 minutes.
- 4) After immersion in ice, 80 µl of 3M potassium acetate/1.8 M formic acid was added and solution left on ice for 30 minutes.
- 5) Following a 15 minute centrifugation in a microfuge (Centrifuge 5414; Eppendorf), the supernatant was transferred to a fresh 1.5 ml tube containing 1 ml of ethanol.
- 6) The solution was mixed by inversion and centrifuged for 20 seconds in a microfuge. The resulting DNA/RNA pellet was drained and resuspended in 100 µl of H₂O.
- 7) 40 µl of 7.4 M ammonium acetate was added to the tube and left on ice for 20 minutes. 0.5 ml of ethanol was added and mixed by inversion. The DNA was precipitated by centrifugation for 10 minutes in a microfuge. The pellet was washed with 70 % ethanol, air-dried until just moist and resuspended in 50 µl of H₂O (approximately 100-200 ng/µl of DNA).

2.3.5. Bacterial DNA (for BACs)

DNA was extracted using the method provided by the supplier of the BAC clones (Research Genetics). A summary of the steps involved are indicated below:

- 1) The cells (1.5 ml) in a 1.5 ml tube were centrifuged for 1 minute in a microfuge and the pellet resuspended in 100 μ l of 50 mM glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0 and placed on ice.
- 2) 200 μ l of 0.2 M NaOH, 1 % SDS was added and the solution mixed by inversion (5-7 times) and placed on ice.
- 3) After the addition of 150 μ l of potassium acetate pH 4.8, the solution was mixed by inversion and centrifuged for 6 minutes in a microfuge.
- 4) The supernatant was transferred to a fresh tube and the DNA was precipitated by the addition of 1 ml of ethanol and pelleted by centrifugation for 6 minutes in a microfuge.
- 5) After rinsing the pellet in 70 % ethanol, the DNA was resuspended in 20 μ l of H₂O.

2.3.6. Plasmid DNA (NME2 cDNA cloned in pCR1000 vector (Invitrogen))

Plasmid DNA was extracted using the Plasmid DNA Midi kit from Qiagen according to the manufacturer's recommended protocol.

- 1) The bacterial pellet was resuspended in 4 ml of P1 buffer (50 mM Tris/HCl, 10 mM EDTA, pH 8.0, 100 μ g/ml RNase A).
- 2) After the addition of 4 ml of buffer P2 (200 mM NaOH, 1 % SDS), the solution was incubated at room temperature for 5 minutes.
- 3) 4 ml of buffer P3 (2.55 M potassium acetate, pH 4.8) was added and the solution was mixed gently. The solution was then centrifuged at 15,000 rpm at 4°C for 30 minutes in a Beckman J2-21 centrifuge.
- 4) The supernatant was added to a Qiagen column (equilibrated with buffer QBT: 750 mM NaCl, 50 mM MOPS, 15 % ethanol, 0.15 % Triton X-100) and allowed to flow through by gravity flow.

- 5) 10 ml of buffer QC (1.0 M NaCl, 50 mM MOPS, 15 % ethanol) was added and allowed to drip through the column. The DNA was eluted with the addition of 5 ml of buffer QF (1.25 M NaCl, 50 mM MOPS, 15 % ethanol) to the column.
- 6) The DNA was precipitated with 0.7 volumes of isopropanol and centrifuged at room temperature for 30 minutes at 1500 rpm in a Beckman GP Centrifuge.
- 7) The DNA was washed with 70 % ethanol, air dried and resuspended in TE pH 8.0, to an approximate concentration of 1 mg/ml.

2.3.7. Elution of NME2 DNA from agarose

The NME2-specific DNA fragment (either after PCR amplification as described in section 2.6.2(1) or after restriction enzyme digestion as described in section 2.4.1) was excised from the gel in a minimum volume of agarose with a clean disposable scalpel and put into a 1.5 ml centrifuge tube. The DNA was purified from the agarose slice using the GeneClean Kit (BIO101 Inc.), according to the manufacturer's recommended protocol:

- 1) Three volumes of Sodium Iodide was added to the DNA/ agarose block and incubated at 55°C for 10 minutes.
- 2) 5 µl of glassmilk suspension was added and the tube placed on ice for 5 minutes.
- 3) After centrifugation in a microfuge for 5 seconds, the supernatant was removed.
- 4) The pellet was washed with 200 µl of New Wash. This was repeated three times.
- 5) 5 µl of H₂O was added and the mixture incubated at 55°C for 5 minutes.
- 6) After centrifugation in a microfuge for 30 seconds, the supernatant containing the DNA was removed.
- 7) The yield of purified DNA was estimated on an agarose gel by comparison with known concentrations of DNA.

2.4. Restriction Enzyme Analysis

Restriction endonucleases were purchased from Gibco-BRL, New England Biolabs and the Boehringer Mannheim. Restriction enzyme digestions were performed using the buffers supplied by the manufacturers and according to their recommendations.

2.4.1 Restriction analysis of Plasmid DNA

1 µg of DNA was digested for 1 to 2 hours at 37°C using 10 to 15 units each of EcoRI and HindIII in a 20 µl reaction. After digestion, the plasmid DNA was electrophoresed as described in section 2.4.3 to distinguish the vector fragments from that of the cloned insert. The 507 bp cloned NME2 insert was excised from the gel as described in section 2.3.7. This DNA fragment was used as a NME2-specific DNA probe as described in section 2.5.2.

2.4.2. Genomic DNA

Genomic samples containing approximately 7-10 µg of DNA were digested for 3 hours using 20-30 units of restriction endonuclease, usually in a 100 µl reaction. After reducing the volume to approximately 20 µl in a Speedvac microconcentrator (Savant), the digested DNA was electrophoresed as described in the following section.

2.4.3. Electrophoresis conditions

After the addition of 3 µl of loading buffer (0.25% Bromophenol blue, 0.25% xylene cyanol FF, 15% Ficoll in water), the digested DNA was electrophoresed through horizontal agarose gels (NBL or Scotlab) submerged in 1 X TAE (40 mM Tris-acetate pH 7.8, 1 mM EDTA pH 8.0) buffer. Genomic DNA was run at 1 to 2 v/cm and plasmid DNA was run at 3 to 5 v/cm. 0.8 % agarose gels made up in 1 X TAE buffer were used for most applications. 2 % agarose gels were used for D17S74 southern hybridisation for optimal allele separation as some of the alleles were small or of similar sizes.

The 1 X TAE running buffer contained 0.5 mg/ml ethidium bromide which allowed the DNA fragments to be visualised using a UVP transilluminator. DNA fragment lengths were determined by using DNA standards of known size, for example λ DNA digested with HindIII. Permanent records of the gels were obtained by photography through an orange membrane using a Polaroid Land camera and Type 57 film.

2.5. Southern Blotting (Southern, 1975)

Southern blotting was used for the analysis of restriction enzyme digested genomic DNA (described in section 2.4.2). This method was used for genotyping individuals from breast and breast-ovarian cancer families using cloned DNA probes (sections 2.2.7 and 3.2). It was also used for the chromosomal localisation of NME2 (section 3.9).

2.5.1. Transfer of DNA from agarose gels

After electrophoresis, the DNA was transferred to a Hybond N+ nylon membrane (Amersham International):

- 1) The gel was soaked in 0.4 M NaOH for 20 minutes with gentle agitation.
- 2) The gel was placed on a wick of 3MM blotting paper which had been soaked in 0.4 M NaOH and draped over a reservoir of the same solution. The nylon membrane which had been pre-soaked in 0.4 M NaOH was placed on top of the gel, followed by a sheet of 3MM blotting paper, also soaked in 0.4 M NaOH. A stack of blotting towels was placed on top of the 3MM paper, followed by a glass plate and a 0.5-1kg weight.
- 3) The denatured DNA was transferred to the membrane by capillary transfer in the NaOH for 3-16 hours.
- 4) After blotting, the membrane was washed in 5 X SSC (from a 20 X stock solution: 3 M NaCl, 0.3 M Tris. sodium citrate pH 7.0) and wrapped in Saranwrap prior to use.

2.5.2. Generation of labelled DNA probes

The randomly primed labelling technique (Feinberg and Vogelstein, 1983) was used to generate radioactively labelled DNA fragments for use as probes in Southern blot analyses. This was performed using the Amersham Multiprime kit according to the manufacturers protocol:

- 1) 5- 50 ng of DNA resuspended in a suitable volume of H₂O in a 1.5 ml tube was denatured by boiling for 5 minutes and placed on ice.

- 2) 10 µl of Oligolabelling Buffer, 5 µl of Primers (random hexamers), 5 µl of [α -³²P] dCTP (Amersham International) and 2 µl of enzyme (klenow fragment of DNA polymerase I) was added to the denatured DNA and incubated at 37°C for 30 minutes.
- 3) After labelling, the unincorporated dNTPs were removed by running the reaction mixture through a Sephadex G50 column (Pharmacia) equilibrated in 3 X SSC.
- 4) The DNA probe was denatured by boiling for 5 minutes immediately prior to adding to the hybridisation mix (described in the following section). For D17S74, the probe was pre-competed (as described in section 2.5.6) before addition to the hybridisation mix.

2.5.3. Hybridisation of labelled DNA probes to DNA immobilised on nylon membranes

The hybridisation solutions used were as follows:

Prehybridisation Solution

5 X SSC

10 X Denhardt's solution (see below)

0.1% (w/v) SDS

40 mM NaPO₄, pH 6.5

0.1 mg/ml sheared herring sperm DNA (Sigma)

100 X Denhardt's Solution

2% (w/v) BSA (Fraction V, Sigma)

2% (w/v) Ficoll (Sigma)

2% (w/v) Polyvinylpyrrolidone (Sigma)

Hybridisation Solution

Prehybridisation solution

10% (w/v) Dextran sulphate

Denatured DNA probe (see section 2.5.2)

- 1) The membrane (produced as described in section 2.5.1) was prehybridised for 2 to 4 hours in 20 ml of prehybridisation solution at 65°C in a glass tube (Hybaid) in a rotating hybridisation oven (Hybaid).
- 2) The prehybridisation mix was then removed and 20 ml of the hybridisation mix containing the denatured DNA probe was added. Hybridisation was performed for between 12-16 hours at 65°C in the rotating oven.
- 3) After hybridisation, the hybridisation mix was removed and the membrane was washed in a metal staining tray with shaking. One 10 minute wash (0.2 X SSC, 0.1 % (w/v) SDS) was performed at room temperature, then one 10 minute wash in a fresh batch of the same solution at 65°C. A higher stringency wash was performed if required (0.1 % SSC, 0.1 % SDS at 65 °C).

NB. Some hybridisations were carried out using the buffer (7 % SDS, 10 % PEG 6000, 0.1 mg/ml sheared herring sperm DNA (Sigma) for both prehybridisation and hybridisation.

2.5.4. Autoradiography

The membrane was wrapped in plastic film (Saranwrap), placed in an x- ray film cassette and exposed to Kodak X-OMAT AR film at -70°C for 1-5 days to visualise the hybridised fragments. The film was developed using a commercial film developing system (Fugi).

2.5.5. Removal of Radiolabelled Probes

Membranes were stripped of radioactive probes by incubation in boiling 0.05% SDS for 30 minutes. The stripped membranes were soaked and stored in 2 X SSC at 4°C until ready for reuse.

2.5.6. Competition hybridisation

250 µg of sonicated human placental DNA was added to the radioactively-labelled probe cMM86 (D17S74) and boiled for 10 minutes. After a 1 hour incubation at 65°C,

the probe was then added to the hybridisation solution. This was used to “compete” out the repetitive sequences within the probe, leaving only unique DNA sequences available for hybridising to the genomic DNA on the membrane.

2.6. DNA Amplification by the Polymerase Chain Reaction

The polymerase chain reaction (PCR), developed by Saiki *et al.* (1988), was used to amplify DNA between two regions of known sequence.

2.6.1. Basic Protocol

The general guidelines for designing each oligonucleotide primer pair was that each primer was about 50% GC rich, non-self complementary and non-complementary to each other. Ideally the oligonucleotides were between 17-24 nucleotides long with a melting temperature of about 60°C (assuming A, T = 2°C; G, C = 4°C). Microsatellite marker oligonucleotide primer sequences were obtained from the literature, for example, Dib *et al* (1996).

PCR was performed in either 25, 50 or 100 µl reactions containing 30-50 ng template DNA, 1 X PCR buffer (Boehringer Mannheim), 20pmol of each primer, 200mM dNTPs (Boehringer Mannheim) and 1 unit of Taq polymerase (Boehringer Mannheim) on a GeneAmp 9600 thermal cycler (Perkin-Elmer Cetus). Reaction mixtures were generally given 35 cycles of 30 seconds at 94°C, 30 seconds at 55°C and 30 seconds at 72°C.

17 µl of the PCR amplification reaction (plus 3 µl of loading buffer: 0.25% (w/v) Bromophenol blue, 0.25% (w/v) xylene cyanol FF, 15% (w/v) Ficoll in water: all obtained from Sigma), was run on a 2 % agarose gel made up in 1 X TBE buffer (0.045 M Tris-borate, 0.001 M EDTA). 0.5 mg/ml ethidium bromide in the 1 X TBE running buffer allowed the DNA fragments to be visualised using a UVP transilluminator. DNA fragment lengths were determined by using DNA standards of known size, for example, ϕ x174 DNA digested with *Hae*III. Permanent records of the

gels were obtained by photography through an orange membrane using a Polaroid Land camera and Type 57 film.

2.6.2. NME2 localisation

1) Generation of human NME2 PCR fragment for cloning

Using the published NME2 cDNA sequence (Stahl *et al*, 1991), oligonucleotide primers were designed. The primer sequences used were primer A: 5' ATC CCT TCT GCT CTC CCA GC 3' (nucleotide position 16-35 bp) and primer B: 5' GAC CCA GTC ATG AGC ACA AG 3'(nucleotide position 522-503 bp). These primers PCR amplified the predicted 507 bp product (data not shown) of the coding sequence from a human cDNA library derived from HeLa cells that had been stress induced by heat shock (gift from Dr. S. Keyes). The PCR amplification conditions were as described in section 2.6.1. The PCR product was gel purified and cloned as described in sections 2.3.7 and 2.9, respectively.

2) Localisation of human NME2 by PCR amplification.

The PCR amplification conditions were as described in 2.6.1., except the annealing temperature for the cycling was at 60°C. The oligonucleotide primer sequences are as indicated in Figure 3.9 in Chapter 3.

2.6.3. PCR amplification products for sequencing EDH17B1 and EDH17B2

Oligonucleotide primer sequences are as indicated in Figure 3.11(b) in Chapter 3 and the combinations of primers used to generate EDH17B1- and EDH17B2-specific PCR amplification products is as described in section 3.12 and Figure 3.11(a). For primer sets A/ C and B/ D, the PCR amplification conditions were as described in section 2.6.1., except the extension time (72°C) for the cycling was 2 minutes. For primers E to M, one of the oligonucleotides from each pair used for PCR amplification was biotinylated at the 5' end and the PCR amplification conditions were as described in section 2.6.1.

2.6.4. Brush 1 localisation

Primer sequences were designed from the published sequence of Brush 1 (Schott *et al*, 1994) and were as follows: primer A: 5' GTT TAA TGT CTC CTA AGC 3' (nucleotide position 965-983 bp) and primer B: 3' CAT CAG TGT AGC CAA GC 3' (nucleotide position 1128-1112 bp). The PCR amplification conditions were as described in section 2.6.1, except the annealing temperature for the cycling was 60°C. The primers amplified the expected 162 bp when human genomic DNA was the template in the PCR amplification reaction.

2.6.5. D13S178E and D13S181E localisation

Primer sequences for D13S178E (clone EST00273) were designed from the available published DNA sequence (accessed via the Genome Database) and were primer A: 5' CCC AGG TGT GCA GCT AAC CT 3' (nucleotide position 32-51 bp) and primer B: 5' CGG TAA GTA TAT TCA TTT CTG 3' (nucleotide position 239-219 bp). Primer sequences for D13S181E (clone EST00297) were as described in Hawthorn and Cowell (1996). The PCR amplification conditions for both gene loci were as described in section 2.6.1.

2.6.6. D13S178E cDNA analysis

The same pair of primers used for the localisation of D13S178E (previous section) were used to PCR amplify from a number of tissue-specific cDNA libraries (Clontech; obtained from L. Rooke, Human Genetic Resources). A 100 µl aliquot from each cDNA library was boiled for 10 minutes (to lyse bacterial cells) before 2 µl was added to each PCR reaction. The PCR amplification conditions were as described in section 2.6.1.

2.6.7. Envoplakin localisation

Oligonucleotide primers were designed from the last predicted exon of the Envoplakin cDNA sequence. This unpublished sequence was obtained from C. Ruhrberg and Dr.

F. Watt, ICRF Keratinocyte Laboratory, Lincoln's Inn Fields, London. The primer sequences used were 5' GCA GGA GGA GTC GAA GCT G 3' (nucleotide position 4937-4955 bp) and 5' GCA GGC GAG GAA GAG CAT C 3' (nucleotide position 5072-5054 bp). The PCR amplification conditions were as described in section 2.6.1, except the annealing temperature for the cycling was at 60°C.

2.7. Detection of alleles at microsatellite marker loci

Three methods of detection were used and are described below:

2.7.1. Ethidium bromide detection

This method is published in Watkins *et al* (1991).

1) 3 µl of loading buffer (0.25% Bromophenol blue, 0.25% xylene cyanol FF, 15% Ficoll in water) was added to 17 µl of each PCR amplification reaction.

2) Products were run on a 6% non-denaturing polyacrylamide gel (Protogel: Flowgen) in 1 X TBE running buffer using a 21cm X 50cm sequencing electrophoresis kit (Bio-Rad) at 50 W. Gel thickness was doubled to 0.8mm for increased rigidity as the gel has to be manipulated in the ethidium bromide detection step (see below). A DNA standard of known size, for example ϕ x174 DNA digested with *Hae*III (Boehringer Mannheim), was also loaded to aid in gel orientation and allele size estimation. The bromophenol blue (in the loading buffer) was used for estimating the gel running times to achieve good separation of different sized PCR amplification products. For example, good allele separation for a microsatellite marker locus with alleles in the range of 150-180 bp was accomplished when the bromophenol blue loading dye was at the bottom of the gel.

3) After gel running, the portion of the gel with the PCR amplification products (usually the middle third) was removed by placing a piece of blotting paper (Whatman) on the gel and then peeling the gel (adhered to the blotting paper) away from the glass plate. The gel was stained for 5 minutes in 0.5 mg/ml ethidium bromide in water and the PCR amplification products visualised using a UVP transilluminator. A permanent

record of the gel was obtained by photography through an orange membrane using a Polaroid Land camera and Type 57 film.

4) The alleles were scored by eye with an arbitrary numbering system. Confirmation of genotyping was performed by loading the samples in a different lane order and/ or pooling samples in the same lane.

2.7.2. Radioactive detection

1) 2 μ l of the PCR reaction was mixed with 3 μ l of loading buffer (95% (w/v) Formamide, 0.05% (w/v) xylene cyanol , 0.05% (w/v) bromophenol blue, 20 mM EDTA pH 8.0), incubated at 95 °C for five minutes and then placed on ice.

2) The PCR amplification products were run on a 6% denaturing sequencing gel (Sequagel, Flowgen) in 1 X TBE running buffer using a 21cm X 50cm sequencing electrophoresis kit (Bio-Rad) at 50 W (or a 44 cm X 50 cm sequencing electrophoresis kit (Bio-Rad) at 100 W). As in section 2.7.1., the position of the bromophenol blue (in the loading buffer) was used for estimating gel running times.

3) The DNA was transferred by blotting in 10 X TBE for 2-3 hours onto Hybond N+ membrane (Amersham International). This was achieved by placing the membrane (previously soaked in 10 X TBE) onto the gel, followed by three pieces of blotting paper (Whatman) and the glass plate of the sequencing apparatus (the one not attached to the gel).

4) The membrane was washed in 0.4 M NaOH for 15 minutes and then in 2 X SSC for 10 minutes. After washing, the membrane was wrapped in Saranwrap until hybridisation.

5) One of the PCR primers was radioactively labelled using a terminal transferase (TT) labelling kit (Gibco) according to the manufacturer's recommended protocol:

A 50 μ l reaction containing: 100 ng oligonucleotide primer, 1 X TT buffer, 3 μ l of [α -³²P] dCTP (Amersham International) and 1 μ l of Terminal Transferase was incubated for 30 minutes at 37°C.

- 6) The end-labelled oligo was hybridised to the membrane for 2 to 4 hours at 42°C in 20 ml of hybridisation solution (7 % SDS, 10 % PEG 6000, 250 mM NaCl, 12 mM phosphate buffer, pH 7.0) in a rotating hybridisation oven (Hybaid).
- 7) After hybridisation, the membrane was washed in a metal staining tray with shaking. One 10 minute wash (2 X SSC, 0.1% (w/v) SDS) was performed at room temperature.
- 8) The membrane was wrapped in plastic film (Saranwrap), placed in an x-ray film cassette and exposed to Kodak X-OMAT AR film at -70°C for 3-12 hours to visualise the hybridised PCR amplification fragments. The film was developed using a commercial film developing system (Fugi).
- 9) The alleles were scored by eye with an arbitrary numbering system. Confirmation of genotyping was performed by loading the samples in a different lane order and/or pooling samples in the same lane.

2.7.3. Fluorescence detection

One of the two oligonucleotide primers used in the PCR amplification reaction was labelled at the 5' end with one of two fluorescent dyes: FAM (blue) or HEX (green). 0.5 µl of each PCR amplification product(s) was mixed with 0.5 µl of a ROX(red)-labelled internal DNA size standard (GS2500 or GS350; Perkin Elmer) and 2 µl of deionised formamide. Prior to loading, the PCR products were denatured for 5 minutes at 95°C and plunged into ice. Samples were run on a 12 cm (well to read) denaturing polyacrylamide gel (Sequagel) at 800 V for 4- 5 hours (depending on allele size) on a 373A Automated DNA Sequencer running Genescan software (Applied Biosystems).

Alleles were analysed using Genotyper software (Applied Biosystems) according to the manufacturer's protocol. As there is an internal DNA size standard in each lane, each allele could be accurately sized in base pairs. Multiplexing of PCR products was also performed by loading microsatellite markers of different allele size ranges and fluorescent dye labelling in the same lane. Alleles were visualised in a graphical

format, with a peak representing an allele. This method was used predominantly in this study after the laboratory acquired the 373A DNA Sequencer and the appropriate computer software (Applied Biosystems).

2.8. Assessment of allele loss

Allele loss was determined as complete or almost complete reduction to hemizyosity in constitutively heterozygous individuals. Complete loss of the deleted allele was rarely evident due to the presence of a small percentage of contaminating normal cells (stroma) in each tumour sample. Loss was scored using the allele detection systems described in sections 2.7.1 and 2.7.2 if there was a clear difference in allele intensity in the tumour profile compared to those for the normal DNA sample. Loss was scored using the allele detection system described in section 2.7.3 if there was a clear difference in allele peak heights in the tumour profile compared to those for the normal DNA sample.

2.9. NME2 cloning

The 507 bp NME2 PCR amplified fragment (as described in section 2.6.2(1)) was gel purified (as described in section 2.3.7) and cloned into the plasmid vector pCRTM1000 using the TA cloning kit (Invitrogen).

2.9.1. Ligation

The ligation reaction was performed in a 10 µl volume according to the manufacturer's recommended protocol with an approximate 1:1 molar ratio of vector to PCR amplification product:

50 ng of vector

25 ng of NME2 PCR amplification product

1 X ligation buffer

1 µl of T4 DNA ligase

The reaction was incubated for 12 hours at 12°C.

2.9.2. Transformation

Transformation of *E.coli* was as recommended by the manufacturer's (Invitrogen) recommended protocol:

- 1) 2 μ l of 0.5 M β -mercaptoethanol was added to the vial of competent cells, INV α F (all supplied in kit). The cells were kept on ice.
- 2) 1 μ l of the ligation reaction (section 2.9.1) was added to the vial of cells and placed on ice for 30 minutes.
- 3) The cells were then given a 1 minute "heat shock" at 42°C, and then plunged into ice for 2 minutes.
- 4) 450 μ l of SOC media (2% Bacto tryptone (Difco), 0.5% Bacto yeast extract (Difco), 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose) was added to the vial and incubated at 37°C in a shaking incubator for 1 hour.
- 5) 10 μ l of the transformation mix was spread onto a 10 cm diameter LB (1 % (w/v) Bacto-tryptone (Difco), 0.5 % (w/v) Bacto-yeast extract (Difco), 1 % (w/v) NaCl, pH 7.0 with NaOH, 1.5 % (w/v) Bacto-agar (Difco)) agar plate containing 50 μ g/ml of kanamycin (Sigma). This antibiotic selected for bacterial cells carrying the plasmid. Prior to plating, 25 μ l of X-Gal (40 mg/ml stock solution: Boehringer Mannheim) was spread over the plate and allowed to diffuse into the agar over a period of one hour. X-Gal allows blue/white colour selection (recombinants are white).
- 6) After an overnight incubation at 37°C, a number of white colonies were selected and the cells grown up in 400 ml of L-broth (as for LB agar but without the Bacto-agar) containing 50 μ g/ml of kanamycin. Plasmid DNA was extracted as described in section 2.3.6. Positive clones were identified by restriction enzyme digestion as described in section 2.4.1.

2.10. Sequence Analysis of DNA

DNA sequencing was performed as described for the Sequenase system (United States Biochemicals), which is a modification of the chain termination protocol (Sanger *et al*, 1977).

2.10.1.Denaturation of Plasmid DNA

This method was used for the preparation of denatured plasmid DNA containing the cloned 507 bp NME2 PCR amplification product.

2 µg of template DNA was denatured in a 10 X volume of 0.2 M NaOH and 0.2 mM EDTA pH 8.0, at 37°C for at least 30 minutes. The DNA was precipitated by the addition of 1/11 th volume of sodium acetate pH 4.5 and 2 volumes of ethanol. After centrifugation the resulting pellet was washed with 70% ethanol, dried briefly and resuspended in 7 µl of H₂O for direct DNA sequencing (section 2.10.3).

2.10.2. Dynabead separation of double stranded DNA

Dynabeads (Dynal) were used to generate single-stranded DNA for sequencing using the manufacturer's recommended protocol. This method was used for the sequence analysis of EDH17B1 and EDH17B2. One of the oligonucleotides in the PCR amplification reaction was biotinylated at the 5' end (section 2.6.3.).

1) 30 µl of Dynabeads were washed in 60 µl of TES (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 100 mM NaCl) in a 1.5 ml tube. After each wash, the tube was placed in a magnet (Dynal) to capture the beads.

2) 40 µl of PCR amplification product was added to the beads and incubated at 28°C for 15 minutes.

3) The tube was placed on the magnet and the supernatant removed. The beads were washed again with TES.

4) 30 µl of 0.15 M NaOH was added, followed by 5 minutes incubation at 28°C .

5) The tube was placed on the magnet and the supernatant removed. The biotinylated DNA strand remained in the tube.

6) The beads/ DNA mixture was washed 5 times with H₂O and then resuspended in 7 µl of H₂O for direct DNA sequencing (section 2.10.3).

2.10.3. Sequencing Reactions

Sequencing reactions were performed using the Sequenase kit according to the manufacturer's recommended protocol. Composition of the solutions used are below the protocol summary.

1) For the annealing reaction, 7 μ l of DNA template (either from sections 2.10.1 or 2.10.2) was mixed with 2 pm of primer and 1 X annealing buffer in a 10 μ l reaction. The oligonucleotide primer sequences for sequencing EDH17B1 and EDH17B2 PCR amplification products are as described in Chapter 3: section 3.12. The oligonucleotides, T7 Promoter and M13 Forward Primer were used to confirm the integrity of the cloned NME2 DNA fragment. These correspond to vector sequences flanking either side of the cloning site.

2) The annealing reaction was incubated at 65⁰C for 2 minutes before cooling to room temperature for 5 minutes.

3) For the labelling reaction 10 μ l of the annealed template/primer solution was mixed with 1 μ l of DTT (0.1M), 2 μ l of labelling mix (5X), 5mCi of [³⁵S] dATP α S (Amersham International) and ~3 units of Sequenase Version 2.0 (2 μ l of an 8-fold dilution in TE pH 7.6). This reaction was incubated at room temperature for 3 minutes.

4) 3.5 μ l of the labelling reaction was added to 2.5 μ l of each of the four dideoxy termination mixes, which had been prewarmed to 37⁰C, in four separate tubes. These mixtures were incubated for 3 minutes, before the reactions were stopped by the addition of 4 μ l of stop solution.

5X Annealing Buffer

200 mM Tris.HCl pH 7.5

100 mM MgCl₂

250 mM NaCl

5X Labelling Mix

7.5 μ M dCTP

7.5 μ M dGTP

7.5 μ M dTTP

ddA Termination Mix

80 μ M dATP

80 μ M dCTP

80 μ M dGTP

80 μ M dTTP

8 μ M ddATP

50 μ M NaCl

ddC Termination Mix

80 μ M dATP

80 μ M dCTP

80 μ M dGTP

80 μ M dTTP

8 μ M ddCTP

50 μ M NaCl

ddG Termination Mix

80 μ M dATP

80 μ M dCTP

80 μ M dGTP

80 μ M dTTP

8 μ M ddGTP

50 μ M NaCl

ddT Termination Mix

80 μ M dATP

80 μ M dCTP

80 μ M dGTP

80 μ M dTTP

8 μ M ddTTP

50 μ M NaCl

Stop Solution

95% (v/v) formamide

20 mM EDTA pH 8.0

0.05% (w/v) Bromophenol Blue

0.05% (w/v) Xylene Cyanol FF

2.10.4. Sequencing Gels

Heat-denatured samples were run on a 6 % denaturing sequencing gel (Protogel, Flowgen) in 1 X TBE running buffer using a 21 cm X 50 cm sequencing electrophoresis kit (Bio-Rad) at 50 W. The wells were formed with a sharks tooth comb. After running, the gel was removed from the glass plate with a piece of blotting paper and fixed in a 1:1 mixture of 10% acetic acid:10% methanol for 10 minutes. The

gel surface was covered in cling film and the gel was dried for 1 to 2 hours (Bio-Rad Model 583 gel dryer, cycle two). Sequencing gels were then autoradiographed and exposed at room temperature overnight or for several days.

2.11. Yeast Artificial Chromosome (YAC) identification

YACs were identified from the whole genome CEPH (Centre d'Etudes du Polymorphisme Humaine) library which contains approximately 33, 000 clones with an average human insert size of 0.9 Mb, representing the equivalent of ten haploid genomes (Cohen *et al*, 1993). This was achieved by two approaches, direct selection and PCR library screening:

2.11.1. Direct selection

To date, 225 YAC contigs spanning an estimated 75 % of the human genome have been constructed (Chumakov *et al* , 1995). The CEPH YAC database was searched using Genethon's Quickmap program suite via the computing facility accessed at the Human Genome Mapping Project (HGMP) Resource Centre, Hinxton Hall, Cambridgeshire, UK. This allowed direct selection of YACs which were likely to contain microsatellite marker loci mapping in the two regions of interest (and YACs which may overlap with these clones) in this study (BRCA2: Chapter 4 and TOC: Chapter 6). Each identified YAC clone (given a grid reference number representing its position in the library) was ordered from the HGMP Resource Centre (section 2.2.8). YACs were propagated and manipulated as described in section 2.3.4.

2.11.2. YAC library screening by PCR amplification

CEPH YAC DNA library pools were obtained from Research Genetics, Huntsville, USA. Microsatellite marker loci (STSs) were screened across the YAC DNA samples by PCR amplification as described in section 2.6.1 in a PCR reaction volume of 25 µl. Presence or absence of STS-specific PCR amplification products was determined by

ethidium bromide detection of DNA on 2% agarose gels (section 2.4.3). The identification of each YAC involved a two tier screening procedure:

1) Primary screening of YACs. 24 block pools were screened by PCR amplification with each block containing 8 pooled microtitre plates of YACs. A positive determined which block of YACs to screen in the next tier.

2) Secondary screening of YACs. One block is divided into 28 pools which are then screened by PCR amplification to give the exact position of the YAC in one of the 8 microtitre plates within that “block”. Three positive PCR amplification products indicate the plate, and the column and row within that plate in which the YAC is located. Any identified YAC clone could then be obtained from the HGMP Resource Centre (or Research Genetics) as in section 2.2.8.

2.11.3. Growth of YACs

1) Plating

Each YAC was streaked on agar (see below) in a petri dish and grown up at an incubation temperature of 30°C for 1- 2 days. An individual clone was “picked” with a sterile pipette tip and inoculated into 5 mls of broth and grown up as described in the following section.

AGAR: 0.8% yeast nitrogen base (Difco), 0.01% adenine HCL, 0.0055% tyrosine, 1.1% casamino acids (Difco) and 2% agar. After autoclaving, 2% membrane sterilized dextrose and 0.005% ampicillin (Sigma) was added.

2) Broth

YACs were grown up at 30 °C in a shaking incubator for 1-2 days in 5 ml of broth (as for the agar described above but without the addition of agar). Cells were then pelleted and DNA extracted (as described in section 2.3.4). A stock of cells for each YAC was stored at -80°C in 20% glycerol.

2.12. BAC library screening by PCR amplification

BAC DNA library plates and clones were obtained from Research Genetics, Huntsville, USA. Microsatellite marker loci (STSs) and Envoplakin mapping oligonucleotides (see section 2.6.7) were screened across the BAC DNA samples by PCR amplification as described in section 2.6.1, in a PCR reaction volume of 25 μ l. Presence or absence of STS-specific PCR amplification products was determined by ethidium bromide detection of DNA on 2 % agarose gels (see section 2.4.3). The identification of each BAC involved a two tier screening procedure:

1) Plate screen of BACs. 24 pools were screened by PCR amplification with each pool containing 384 BAC clones. A positive determined which plate of BACs to screen in the next tier.

2) Individual BAC clone screen. PCR screening of the 384 BACs within each positive plate was performed as follows:

a) A 384-pin device (Genetix) was used to spot the clones onto a LB agar plate with 12.5 μ g/ml of chloramphenicol (Sigma) added to select for bacterial cells containing the vector. These clones were grown up overnight in a 37°C incubator.

b) A small amount of each clone was directly added to a 25 μ l PCR reaction (as described in section 2.6.1). A pooling strategy was used to minimise the number of PCR reactions required before identification of the individual positive BAC clone. A small amount of the positive BAC clone was used to inoculate 5 ml of L-Broth with 12.5 μ g/ml of chloramphenicol added and grown up overnight in a 37°C shaking incubator. DNA was extracted as described in section 2.3.5.

2.13. Linkage analysis

Linkage analyses for the breast cancer study was performed by Prof. T. Bishop and G. Crockford (both from ICRF Genetic Epidemiology Unit). S. Bryant (ICRF Human Genetic Resources) performed the analyses for the NEPPK study. The models used in the analyses are briefly described below:

2.13.1. BRCA1

2.13.1.1. Two-point linkage analysis

Lod score analysis was performed using the MLINK program of LINKAGE version 5.2 and FASTLINK version 2.3P (Lathrop *et al*, 1988; Schaffer *et al*, 1994). Statistical analysis was conducted based on the CASH model derived by Claus *et al*, (1991), modified as described in Easton *et al*, (1993b). Breast cancer susceptibility was modelled as an autosomal dominant allele, with a population frequency of 0.0033. Age-specific liability classes were incorporated into the analysis accumulating to an 80 % breast cancer risk at age 70.

2.13.1.2. Multipoint linkage analysis

Statistical analysis was performed as described in the previous section. Multipoint analysis was conducted using microsatellite marker loci proximal (either D17S250, THRA1, D17S857, D17S800 or D17S776), intragenic (D17S855) and distal (either D17S579 or D17S791) to BRCA1. Allele loss was incorporated into the analysis based on a model proposed by Dr. D. Easton and Prof. T. Bishop (personal communication). Basically, if allele loss was detected in a tumour, the affected women displaying loss was replaced in the pedigree by a person of unknown phenotype with the same genotype information. This individual was then given an imaginary offspring whose affected status was identical to the original individual replaced in the analysis. This offspring had a genotype that gave unequivocal evidence that the retained alleles from the allele loss analysis were transmitted to this individual.

2.13.2. NEPPK

Linkage analysis was performed with the MLINK module of the FASTLINK software package version 2.2 (Cottingham *et al*, 1993), assuming an autosomal dominant inheritance for NEPPK, with a gene frequency of 0.003 and a penetrance of 100%.

Allele frequencies for the microsatellite marker loci were either computed from the data or else were estimated to be equifrequent.

2.14. Estimation of allele frequencies

The frequency of the disease allele shared at the D17S1603 locus for the focal NEPPK/ oesophageal families from the UK and Germany was estimated as follows. CEPH DNA samples (section 2.2.9) were used as templates in PCR amplification reactions using D17S1603 primers as described in section 2.6.1. Alleles were detected and sized as described in section 2.7.3. In total 33 of these ‘random’ individuals were genotyped at this locus. The allele frequencies were calculated as a fraction of the number of chromosomes with each allele over the total number of chromosomes analysed (66 in total).

CHAPTER 3: GENETIC MAPPING OF BRCA1

3.1. Introduction

In this chapter, genetic analysis of the long arm of chromosome 17 in sixteen breast and breast-ovarian cancer families was conducted to confirm linkage to the putative breast cancer predisposition locus, BRCA1, and to identify any recombinants which may reduce the region in which this gene lies. The focus for this study was a large breast-ovarian cancer family designated BOV3. Analysis and exclusion of a number of candidate genes for BRCA1 was performed.

3.2. Linkage analysis of 16 breast and breast-ovarian cancer families

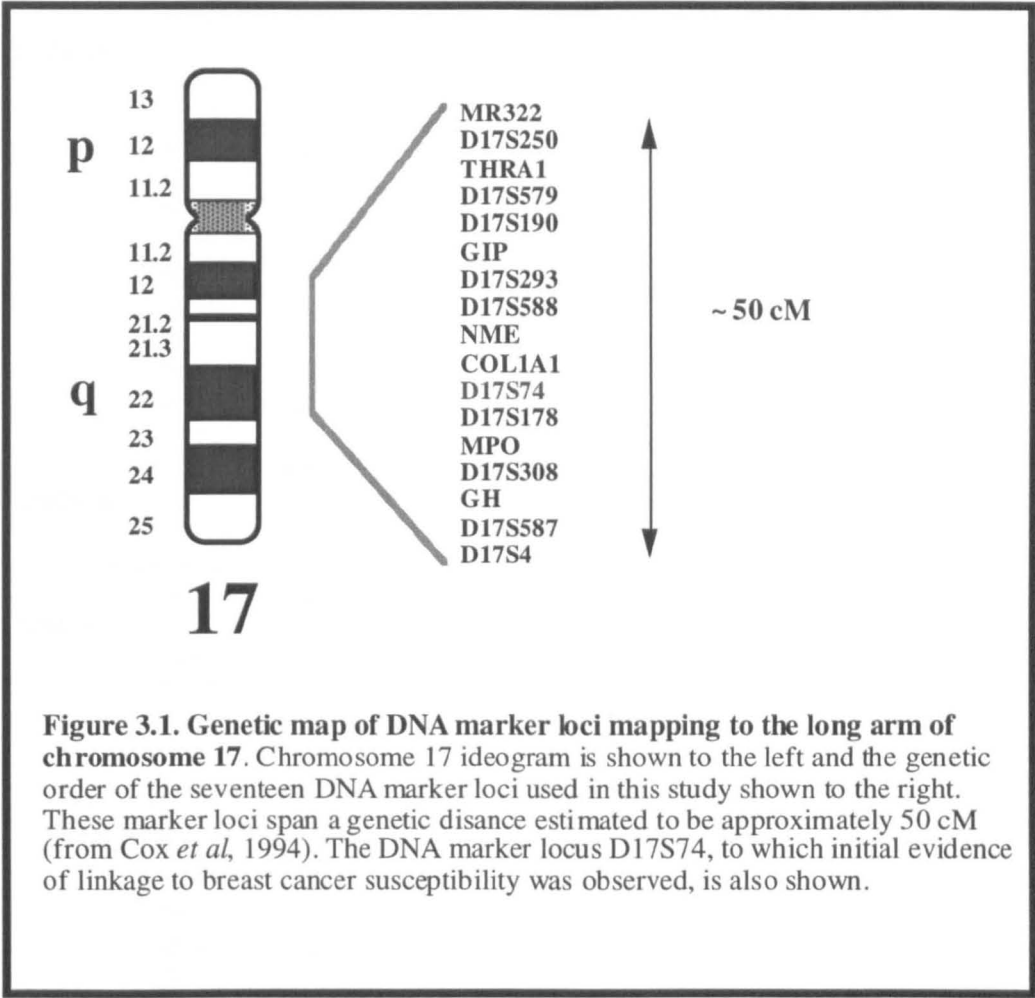
17q linkage analysis was conducted on 16 families, selected from a set of 180 ascertained as described in section 2.1.1. These families were chosen as being the most informative for linkage studies by the following criteria:

- (i) consistency with dominant inheritance for early onset breast or breast- ovarian cancer
- (ii) availability of DNA samples from family members, in particular, affected individuals

Thirteen of these pedigrees were classified as site-specific breast cancer families due to the presence of breast cancer in noticeable excess over the expected incidence of sporadic occurrence. The remaining three families were classified as breast-ovarian cancer, due to the presence of one or more women who are at high risk of being gene carriers (25% or more, as determined by pedigree analysis) who had developed ovarian cancer.

Lymphocyte DNA samples (see sections 2.2.1, 2.3.1 and 2.3.2) from informative family members were genotyped (as described in sections 2.5, 2.7.1 and 2.7.2) from a selection of 17 DNA marker loci mapping to the long arm of chromosome 17 (Figure 3.1). These loci span a sex-averaged genetic distance of approximately 50 cM (Cox *et al*, 1994). Statistical analysis was performed by Prof. T. Bishop and G. Crockford

following the model described in section 2.13.1.1. Two-point LOD scores for some of the DNA marker loci are shown as an aggregate for all families analysed and, for the more informative markers, by family in Table 3.1.



The Variable Number Tandem Repeat (VNTR) DNA marker locus, D17S74, originally indicating linkage to BRCA1 (Hall *et al*, 1990), provided weak evidence for linkage to the disease in these families with a maximum aggregate LOD score of 0.92 at a recombination fraction of 0.20. However, the LOD scores for some of the DNA marker loci proximal to D17S74 suggested that a breast cancer susceptibility gene in a number of the ICRF families may be located more centromeric than D17S74. The strongest evidence for linkage to the disease was from the large breast/ovarian pedigree, BOV3, with the highest LOD score of 2.35 at a recombination fraction of 0.001 at the NME1 locus. The magnitude of this LOD score implied that cancer in this

Table 3.1. Two-point LOD scores for 17q DNA marker loci segregating in 16 breast cancer families

Marker	Family	LOD score at recombination fraction of:						
		.001	.01	.05	.10	.20	.30	.40
THRA1	BOV3	.59	.58	.52	.45	.31	.19	.09
D17S579	BOV3	1.58	1.72	1.89	1.83	1.45	.93	.37
	BC19	-.22	-.20	-.14	-.10	-.04	-.01	.00
	BC1	-.14	-.13	-.09	-.06	-.02	.00	.00
	BC2	-.23	-.21	-.17	-.13	-.07	-.03	-.01
	BC3	.23	.23	.21	.18	.12	.06	.02
	BC9	.09	.09	.07	.05	.03	.01	.00
	BC12	.15	.15	.13	.11	.07	.03	.01
	BC14	.21	.20	.17	.14	.08	.03	.01
	BC15	.19	.19	.16	.13	.08	.04	.01
	BC20	-.60	-.55	-.41	-.28	-.14	-.06	-.01
	BC22	.07	.07	.06	.05	.03	.01	.00
	BC23	.22	.21	.16	.12	.05	.01	-.01
	TOTAL	1.58*	1.77	2.07*	2.05*	1.64	1.03*	.38*
D17S588	BOV3	2.11	2.06	1.88	1.65	1.14	.60	.14
	other families	.16	.23	.41	.48	.40	.23	.06
	TOTAL	2.27*	2.29	2.29	2.13	1.54	.83	.20
NME1	BOV3	2.35	2.30	2.09	1.82	1.27	.71	.25
	BC19	.34	.33	.28	.22	.13	.06	.01
	BC1	-.14	-.13	-.10	-.06	-.02	.00	.00
	BC2	.21	.20	.18	.15	.09	.04	.01
	BC3	.57	.55	.49	.41	.26	.13	.03
	BC7	-.37	-.35	-.27	-.20	-.10	-.04	-.01
	BC12	-.15	-.15	-.13	-.10	-.06	-.03	-.01
	BC15	.13	.13	.11	.09	.05	.02	.01
	BC20	.46	.45	.40	.34	.23	.11	.03
	BC23	.02	.02	.02	.01	.01	.00	.00
	BOV2	.20	.19	.16	.13	.08	.04	.01
	TOTAL	3.63*	3.56*	3.25*	2.81	1.94	1.04	.29*
D17S74	BOV3	-.15	.35	1.03	1.26	1.19	.85	.39
	other families	-1.15	-1.08	-.83	-.58	-.27	-.13	-.07
	TOTAL	-1.30	-.73	.20	.68	.92	.72	.32

* entries do not exactly sum to "total" because of rounding error.

family is highly likely to be due to BRCA1. NME1 maps centromeric to D17S74, with a genetic distance of 3cM in males and 11cM in females (Easton *et al*, 1993b). The highest aggregate LOD score was also obtained at the NME1 locus (3.63 at $\theta = 0.001$). Although the majority of evidence came from the BOV3 family, five other families gave small positive LOD scores at NME1: BC19, BC3, BOV2, BC2 and BC20 with maximum LOD scores of 0.34, 0.57, 0.20, 0.21 and 0.46 at a recombination fraction of 0.001, respectively. Multilocus haplotypes were constructed for the BOV3 family and the latter five families in an effort to define more precisely the region in which BRCA1 may lie. The most informative of these families are discussed below (Figures 3.2, 3.3).

3.2.(a) BOV3

Haplotype data for members of the BOV3 family is shown in Figure 3.2. A number of recombination events with the disease and the DNA marker locus, MR322, suggested that the disease gene lay distal to this centromeric marker, with affected individuals either bearing a "1" allele (3540, 3563, 3550 and 3566) or a "2" allele (3677 and 3684) or a "3" allele (3543) on the disease chromosome. The microsatellite marker locus, D17S250 which maps distal to MR322, revealed no recombination with the disease. For example, D17S250 had a "1" allele shared by all affected individuals tested which suggested that the recombination event had been resolved proximal to this locus. However, D17S250 was homozygous for the "1" allele in a number of individuals which could mask potential recombination events. The telomeric boundary for BRCA1 appears proximal to D17S74 which shows clear recombination events within the family: 3563 has the "5" allele and the mother-daughter pair (3677 and 3684) share the "1" allele whilst the affected individuals (3540, 3550, 3566, 4351 and 4352) share a "6" allele at this locus on the disease chromosome. These cross-overs appear to be resolved at the NME1 locus with all typed affecteds bearing a "1" allele on the putative disease chromosome. The ovarian cancer case 3552 shared the disease haplotype distal to the marker D17S588 which suggested that BRCA1 lay telomeric to this marker locus. Her mother (3540) had breast cancer diagnosed at age 42 and her affected sister


Figure 3.2. Haplotype construction for 17q DNA marker loci in BOV3


Affected status is indicated from the key in left hand corner of the figure with age of diagnosis (dx) underneath the affected individual (if known). Sex of individuals is as described in Figure 1.1. A diagonal slash indicates individual deceased. The order of DNA marker loci is shown in the top left corner of the figure. Haplotypes are displayed under the individual. Numbering of alleles for each marker locus is arbitrary. The putative disease haplotype segregating through the family is indicated in red text. Inferred haplotypes were constructed from genotyped family members for 4345, 3535, 4351, 4352, 3541 and 3552. Haplotype construction for 4351, 4352 and 3552 was inferred from the typing of spouses and offspring; this genotyping is not shown as it is clinically relevant. A zero (0) allele indicates allele unknown due to not being typed, failed PCR reaction or allele status cannot be inferred.

MARKERS:

MR322
D17S250
THRA1
D17S190
D17S579
GIP
D17S293
D17S588
NME1
D17S74
D17S178
GH
D17S587
D17S4

KEY:

 Breast cancer

 Ovarian cancer

 Other cancer

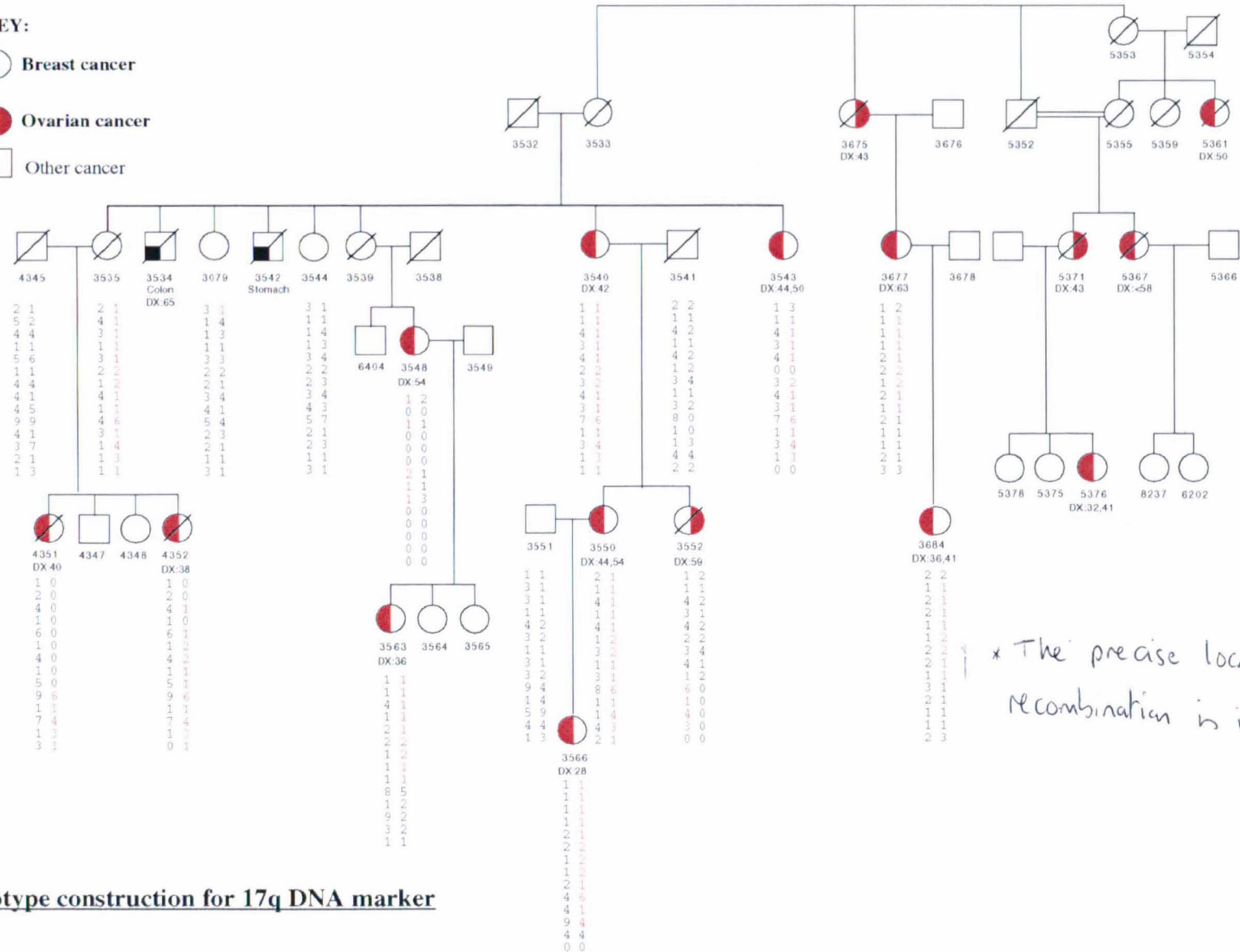


Figure 3.2. Haplotype construction for 17q DNA marker loci in BOV3

(3550) had bilateral breast cancer at ages, 44 and 54. This would suggest that 3552 was a BRCA1 gene carrier. If this were the case, BRCA1 would be localised in the chromosomal interval spanned by D17S588 and D17S74 from haplotype analysis of the BOV3 family. The other possibility was that 3552 had sporadic disease, particularly as the ovarian cancer was diagnosed at 59 years, which is a relatively late age of onset for familial cancer. Further studies on this individual and other members of this family are described (section 3.5).

3.2.(b) BOV2

BOV2 consisted of only two affected sisters; one individual 489 had both breast and ovarian cancer (Figure 3.3a). The affected sister pair demonstrated clear discordance for D17S250 but not for the more distal marker loci. This suggested BRCA1 lay telomeric to D17S250, but the limited size of this family prohibited statistical confirmation of linkage to BRCA1.

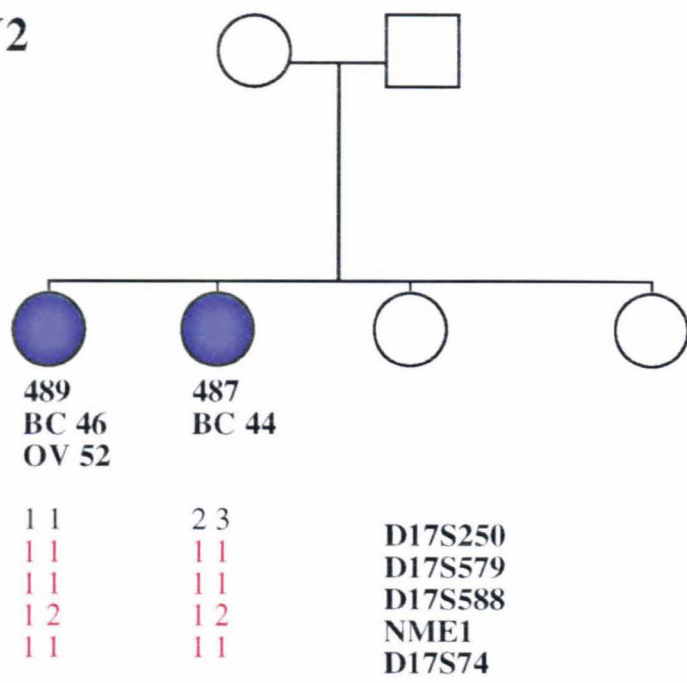
3.2.(c) BC19

BC19 was a site specific breast cancer family with three affected women. A consistent disease haplotype was shared distal to the marker D17S579, with clear discordance at this locus and centromeric to it (Figure 3.3b). This suggested, if the disease was linked to BRCA1 in this family, that BRCA1 maps telomeric to D17S579.

Figure 3.3. Haplotype construction for 17q DNA marker loci in families: (a)BOV2 and (b) BC19.

Affected status and age of diagnosis indicated under the affected individual. BC=breast cancer, OV=ovarian cancer. Sex of individuals is as described in Figure 1.1. The order of DNA marker loci are shown to the right of each pedigree. Haplotypes are displayed under each typed individual. The putative disease haplotype in BC19 is indicated in red. For BOV2, both affected sisters have common haplotypes distal to D17S250 for each chromosome homologue. Haplotype information for the unaffected sisters of 489 and 487 in BOV2 are not shown as they are clinically relevant.

(a) BOV2



(b) BC19

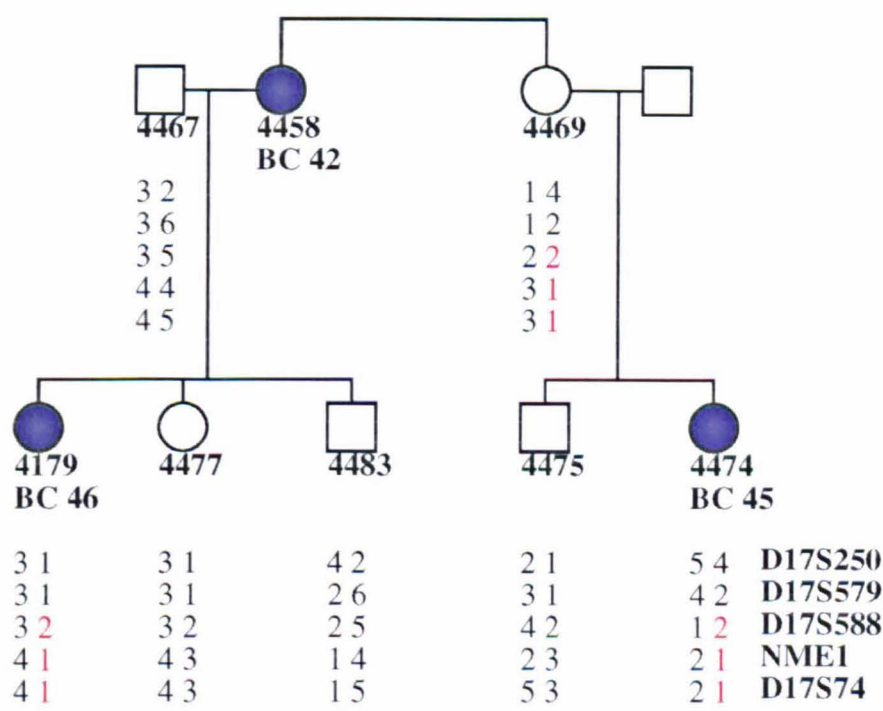


Figure 3.3. Haplotype construction for 17q DNA marker loci in families: (a)BOV2 and (b) BC19.

3.3. Breast Cancer Linkage Consortium Analysis (Easton *et al.* 1993b)

The results of the ICRF 17q linkage data were combined with data from a number of groups worldwide in a collaborative study. The aims of this study were threefold:

- (i) to confirm linkage of breast cancer susceptibility to BRCA1 and to define its genetic location more precisely. This was achieved by each group typing the same set of marker loci centromeric and telomeric to the marker D17S74. Data was submitted without bias towards family structure or the likelihood of a family being linked.
- (ii) to estimate the proportion of BRCA1-linked breast cancer pedigrees.
- (iii) to estimate the penetrance of BRCA1 and possible risk of other malignancies in gene carriers.

214 breast cancer families, including 57 breast/ ovarian families, were typed with six DNA marker loci spanning a region of 30cM on chromosome 17q. The statistical model is summarised in section 2.13.1 (further details in Easton *et al.* 1993b). The DNA marker loci (D17S250, D17S579, D17S588, NME1, D17S74, GH) used in this study are shown in Figure 3.1. Results from the multipoint linkage analysis indicated that the most likely location for the BRCA1 gene was distal to D17S250 and proximal to D17S588; a genetic distance of 8.3cM in males and 18cM in females. The location of the gene within this interval could not be clearly defined due to individual recombination events placing BRCA1 either side of the marker D17S579. Heterogeneity analysis implied that the majority of breast-ovarian cancer families were likely to be due to an inherited mutation in BRCA1. There was clear heterogeneity in breast cancer only families with approximately 45% likely to be due to BRCA1. Breast cancer risk in BRCA1 gene carriers was estimated to be 59% by age 50 years and 82% by age 70 years.

3.4. Re-evaluation of ICRF mapping study

The mapping results generated from the three ICRF families (BOV3, BOV2 and BC19) were re-evaluated with respect to the Consortium study and are discussed

below. An additional 47 ICRF families were also assessed for BRCA1 linkage and some of this work is described in Chapters 4 and 5.

3.4.(a) BOV3

Cancer in the large breast-ovarian family BOV3 is clearly linked to BRCA1 but appeared to provide very little information as to fine genetic localisation of BRCA1. A re-evaluation of the ovarian cancer case 3552 (Figure 3.2) suggests this is likely to be a case of sporadic disease. A number of recombination events with the disease in the consortium data set placed BRCA1 centromeric to D17S588, however, 3552 shared a disease haplotype distal to this marker locus. This family has been genotyped with further marker loci mapping within the minimally defined genetic interval for BRCA1 and the analysis extended to other family members (section 3.5).

3.4.(b) BOV2

From the consortium study, disease in the majority of breast-ovarian cancer families is likely to be BRCA1-linked. Therefore, although BOV2 is only an affected sib-pair, cancer susceptibility in this family is possibly due to BRCA1. Mapping data from this family placed BRCA1 telomeric to D17S250 which is consistent with the consortium data.

3.4. (c) BC19

Family BC19 localised BRCA1 distal to D17S579 but there was no convincing evidence that breast cancer susceptibility in this family is due to BRCA1, particularly as Chamberlain *et al* (1993) provided strong evidence that BRCA1 maps proximal to this marker locus. As BC19 is a breast cancer only family, there is a 55% chance (from Easton *et al*, 1993b) that disease in this family may not be due to an inherited mutation in BRCA1. Recently, individual 4477 (Figure 3.3(b)) was diagnosed with breast cancer at age 55. As this individual does not share a common haplotype with her affected cousin (4474), this further supports the likelihood that disease in this family is unlinked to BRCA1.

3.5. Further genetic analysis of BRCA1 in BOV3

BOV3 is a single extended family with thirteen cases of breast cancer (with ages of diagnosis ranging from 28 to 63 years) and three cases of ovarian cancer (diagnosed at 43, 59 and less than 58 years). Cancer susceptibility in this family is linked to BRCA1 (Table 3.1, section 3.2.a). In this section, genotyping of further individuals is described. Individuals 3552 and 3542 were genotyped from archival pathological material (extracted as described in section 2.3.3). All available individual DNA samples have been genotyped with DNA marker loci mapping within the minimal region defined by the consortium study: D17S250 to D17S588. The majority of these marker loci were highly informative 'CA' microsatellite repeats which were genotyped by PCR amplification (as described in section 2.7.1 and 2.7.2). In total, 29 DNA marker loci have been screened across this family and haplotypes for nine microsatellite marker loci mapping within the minimal region for BRCA1 are shown in Figure 3.4.

3.5.(a). Genetic analysis of the ovarian cancer case 3552

Further mapping of the ovarian cancer case, 3552, confirms that disease in this individual is sporadic. Genotyping of archival DNA from this individual and lymphocyte DNA from the close relatives (3553, 3568, 3567) clearly demonstrated that 3552 had inherited the maternal wildtype BRCA1 chromosome (Figure 3.4). In addition, the fine scale mapping of this region reduces the possibility of an unusual double recombination event having occurred in this individual. Therefore, ovarian cancer in individual 3552 was likely to have occurred sporadically, or as a result of the inheritance of a distinct cancer susceptibility mutation to that segregating in the rest of the family.

3.5.(b). Reduction of the minimal genetic interval for BRCA1

Haplotypes constructed for the affected individual 3684 clearly show maternal inheritance of the disease-bearing haplotype distal to D17S800 and the non-disease

Figure 3.4. Haplotype construction of 17q DNA marker loci mapping between D17S250 and D17S293 in the BOV3 family

Symbols are as defined in Figures 1.1 and 3.2. Age of diagnosis (dx) is shown under the affected individual, if known. The order of DNA marker loci and the putative disease haplotype are shown in the top left hand corner of the figure. The putative disease haplotype is indicated in red. Genotyping for individuals: 4351, 4352, 3535, 3541, 3532 and 3533 are inferred. Some haplotypes are not shown as they are of possible clinical relevance. Individuals 3552 and 5367 were typed from archival breast tumour tissue sections. A zero allele indicates allele unknown due to not being typed, failed PCR reaction or allele status cannot be inferred.

MARKERS: **disease haplotype**

D17S250 1
 THRA1 1
 D17S857 4
 D17S800 1
 GAS/D17S846 1
 D17S855 1
 D17S183 1
 D17S579 1
 D17S293 2

KEY:

Breast cancer



Ovarian cancer

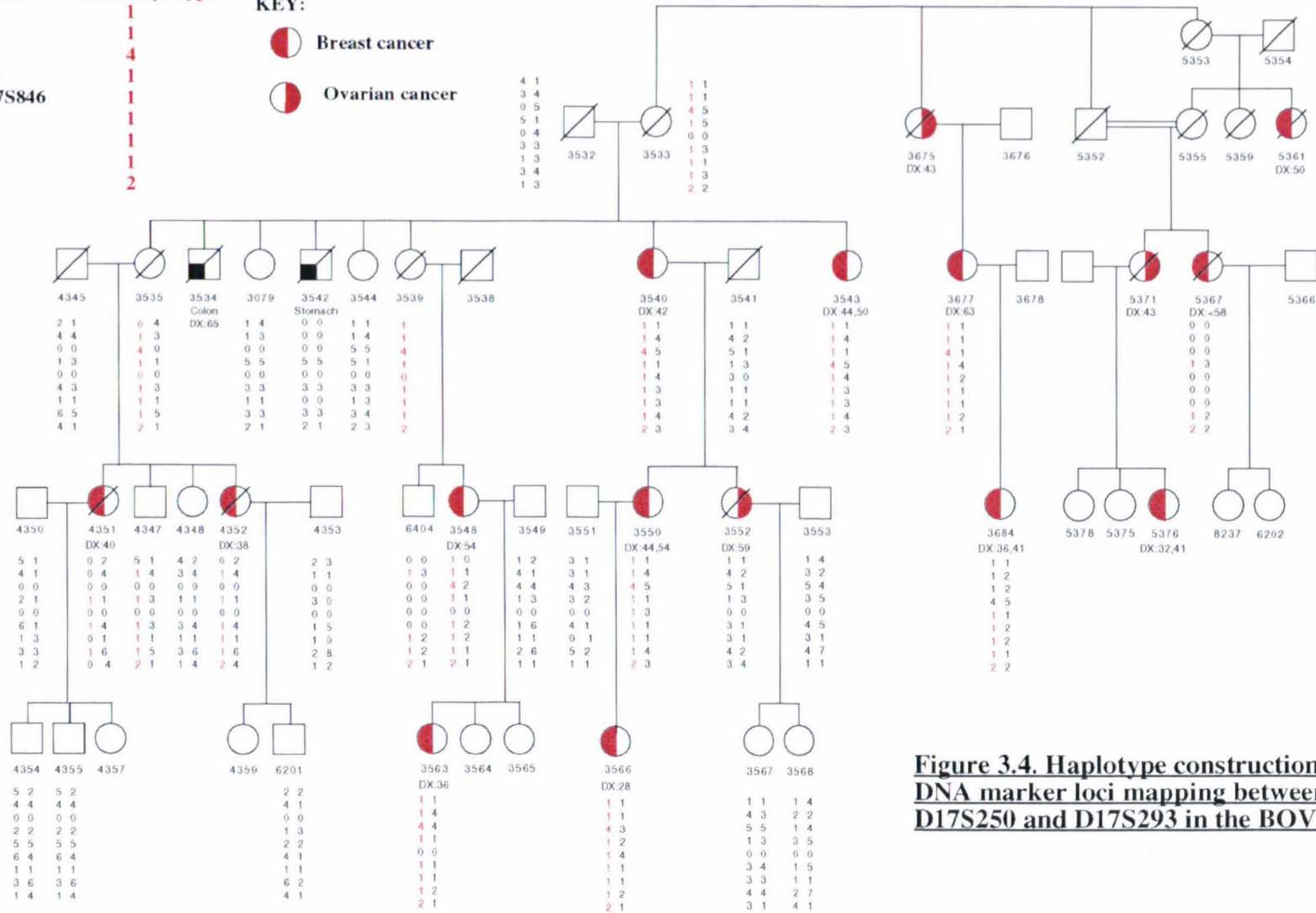


Figure 3.4. Haplotype construction of 17q DNA marker loci mapping between D17S250 and D17S293 in the BOV3 family

Figure 3.5. Mapping the recombination event in individual 3684 from BOV3

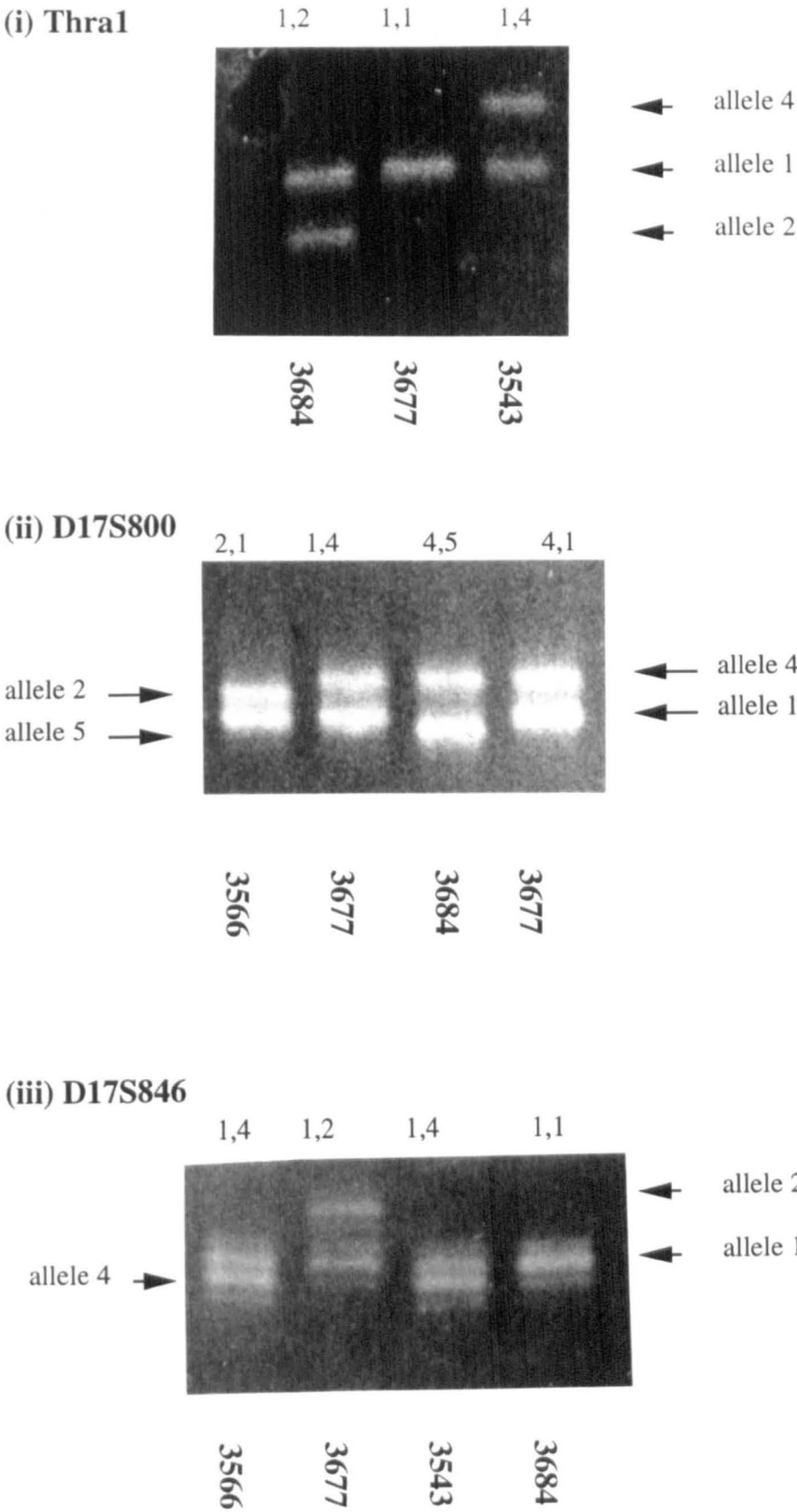
Examples of genotypes scored for individuals 3684, 3677 and other putative BRCA1 gene carriers from the BOV3 pedigree are shown (see Figure 3.4). The alleles for each of the three DNA marker loci shown were detected (as described in 2.7.1) and numbered arbitrarily. The individual identification number is shown under each lane. The number designated for each allele is shown either side of the gel photograph. The genotype is shown above each lane.

(i) **Thra1**. Allele 1 is segregating with the disease in the BOV3 family. 3684 is heterozygous at this locus (1: 2) and has inherited allele 1 from her mother, 3677. However, as 3677 is homozygous for allele 1, the homologue passed onto her daughter cannot be determined.

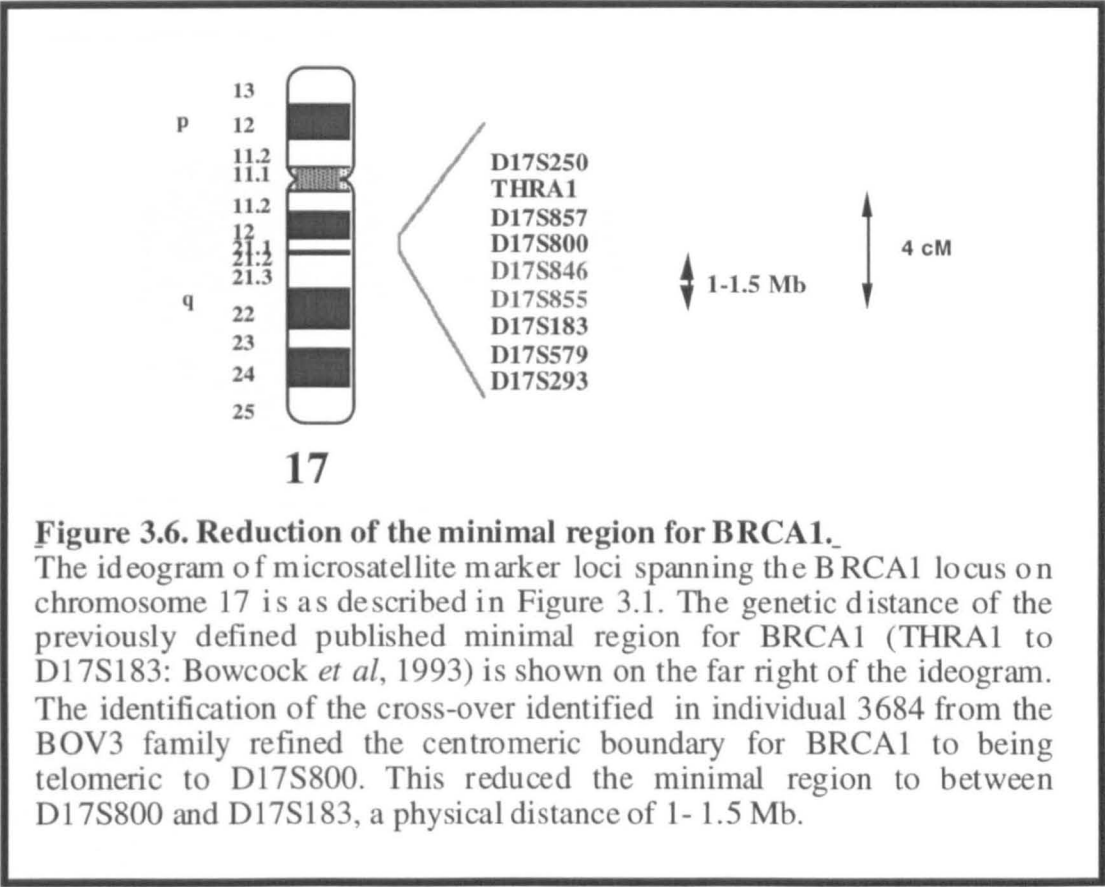
(ii) **D17S800**. Allele 1 is segregating with the disease in the BOV3 family. 3684 has inherited the 4 allele from her mother, 3677, indicating a cross-over with this DNA marker locus and BRCA1.

(iii) **D17S846**. Allele 1 is segregating with the disease in the BOV3 family. 3684 has inherited the 1 allele from her mother 3677 who is heterozygous at this locus (1:2). This indicates the cross-over identified at D17S857 and D17S800 has been resolved at D17S846. DNA marker loci mapping distal to D17S846 confirm this result and demonstrate 3684 has inherited the disease chromosome telomeric to D17S800 (Figure 3.4).

Figure 3.5. Mapping the recombination event in individual 3684 from BOV3



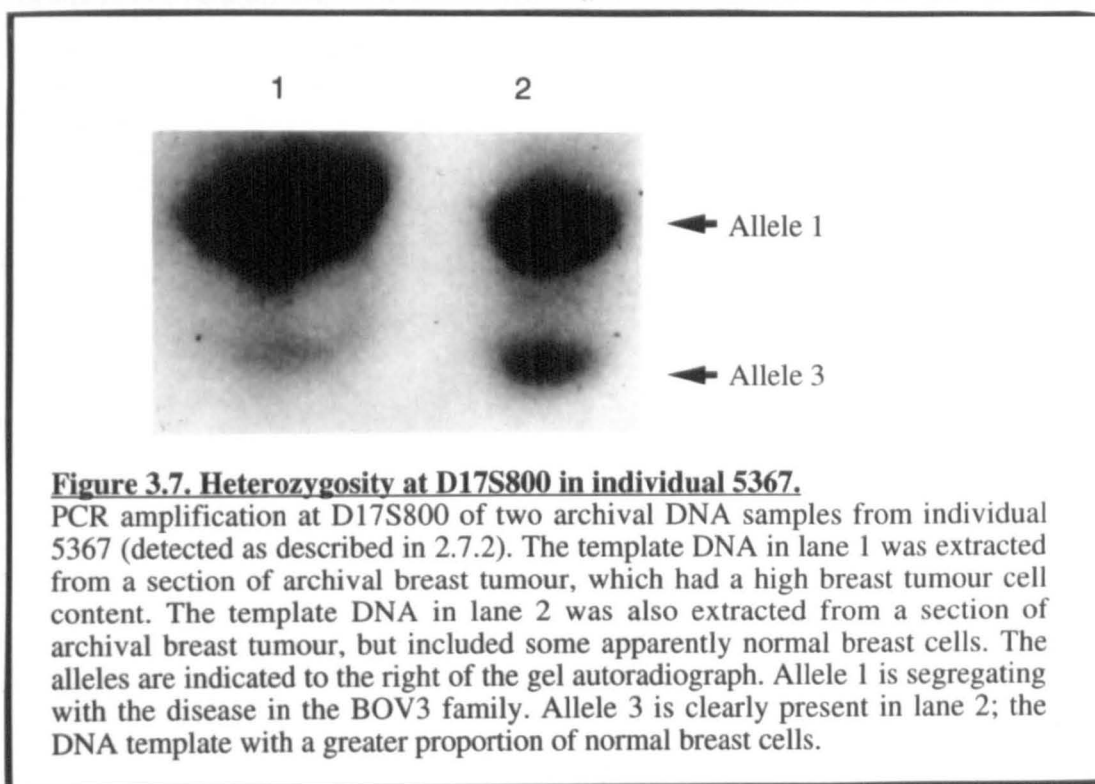
chromosome proximal to D17S846. Fine mapping of this recombination event is shown in Figures 3.4 and 3.5. This cross-over event was unidentifiable from previous haplotype analysis as it was masked by the homozygosity of the two genotyped proximal microsatellite marker loci, D17S250 and THRA1 in the mother (3677) of this individual (Figure 3.5(i)). Individual 3684 had bilateral breast cancer diagnosed at 36 and 41 years, whilst her mother had breast cancer and her grandmother had ovarian cancer. Bilateral disease is a feature of inherited cancer predisposition, particularly if presented at an early age. In addition, individuals 3543 and 5376 also present with bilateral breast cancer in this family. This suggests that 3684 is unlikely to be a sporadic case and that BRCA1 lies distal to D17S800 (Figure 3.6). No other recombination events refining the localisation for BRCA1 were detected in this family.



3.5.(c). BRCA1 homozygote?

A further, intriguing, branch of this family was analysed. This involved the affected daughters of the uncle-niece marriage of 5352 and 5355, who were potentially homozygous for the BRCA1 mutation since both parents were putative carriers (from

pedigree analysis). If this was the case, it would suggest that a fully functional copy of the BRCA1 gene was not required for normal development. However, analysis of DNA from a section of archival pathological breast tissue containing both tumour and apparently normal cells (extracted as described in section 2.3.3) from individual 5367 revealed that this individual was heterozygous for two of the three microsatellite marker loci analysed, D17S800 and D17S579 (Figures 3.4 and 3.7). DNA was unavailable from the parents of 5367 for determination of transmission of the disease chromosome, but it is possible that it was inherited from the unaffected mother 5355 who had a sister (5361), diagnosed with breast cancer at the age of 50. Subsequent to the cloning of BRCA1, putative BRCA1 homozygotes have been identified in two UK families and are discussed in section 5.2.



3.5.(d). Risk of cancer in male BRCA1 carriers

Two males were diagnosed with cancer in this family (3534, 3542) and are at 50% risk (from pedigree analysis) of being carriers of the BRCA1 mutation segregating through the BOV3 family. No material was available for the colon cancer case, 3534, but haplotype analysis of archival pathological DNA from the stomach cancer case,

3542, showed that the individual had not inherited the disease chromosome (Figure 3.4). The Breast Cancer Linkage Consortium analysis of risk of other cancers in BRCA1 gene carriers revealed an increased risk of colon cancer (relative risk [RR] of 4.11) and prostate cancer (RR of 3.33) (Ford *et al*, 1994).

3.5.(e). Summary of genetic mapping in BOV3

The results of the analysis of the BOV3 family confirm cancer predisposition in this family is due to an inherited mutation in BRCA1. The recombination event in individual 3684 places BRCA1 distal to D17S800. A recombination event identified by Bowcock *et al.*, (1993) suggests that BRCA1 is proximal to D17S183. These two recombination events define a minimal region bearing BRCA1 to between D17S800 and D17S183 (Figure 3.6), a distance estimated to be between 1.0 and 1.5 Mb based on a physical map of the region (Brown *et al*, 1995). Analysis of other ICRF breast cancer families revealed no further mapping information but identified other families in which cancer predisposition was due or likely to be due to an inherited mutation in BRCA1 (Chapter 5).

3.6. Candidate gene analysis for BRCA1

Having defined the genetic boundaries flanking BRCA1, genes situated within this chromosomal location could be analysed as potential candidates for BRCA1. In this section, the analysis of a number of candidate genes for BRCA1 are described in historical order. The identification of a locus on 17q predisposing to breast and ovarian cancer suggested a number of potential candidate genes, which map to this chromosome arm, for BRCA1. The quickest and most convincing method of exclusion is identifying affected individuals from BRCA1-linked pedigrees displaying recombination events between the candidate gene and the disease locus. If no recombination events are detected, the gene maps within the minimum region and then mutation studies are performed to identify any potentially disease-causing mutations in BRCA1-linked families. BRCA1 candidate gene analyses are described below.

3.7. NME 1 and 2 (NM23)

NM23 is a heterodimeric protein which acts as a nucleoside diphosphate (NDP) kinase (Gilles *et al*, 1991). NDP kinases are involved in the synthesis of nucleoside triphosphates and the NM23 protein may act in the regulation of signal transduction by complexing with G-proteins causing activation/ inactivation of developmental pathways (Wallet *et al*, 1990; Biggs *et al*, 1990).

Two cDNAs, NME1 and NME2, were isolated from murine K-1735 melanoma cell lines using differential hybridisation (Steeg *et al*, 1988; Stahl *et al*, 1991). Comparison of cell lines displaying low and high metastatic rates indicated two cDNAs that were down regulated in the highly metastatic melanoma cell line. Upon NME1 cDNA (isolated from a cell line showing reduced metastatic potential) transfection into these cells, metastatic potential was significantly reduced. These two NME cDNAs, NME1 (Steeg *et al*, 1988) and NME2 (Stahl *et al*, 1991), have been shown to encode the A and B polypeptide chains of an NDP kinase (NM23), respectively. Both chains consist of 152 amino acid residues and share 88% nucleotide homology. Human NME2 is identical to the B subunit of human erythrocyte NDP kinase (Gilles *et al*, 1991). It has been proposed that NM23 may have a role as a tumour suppressor gene (Steeg *et al*, 1988; Rosengard *et al*, 1989; Steeg and Liotta, 1990).

The human NME1 gene has been assigned to 17q21-q22 (Varesco *et al*, 1992) and could be considered as a candidate for BRCA1. The following section describes the localisation of NME2.

3.8. Cloning of NME2

From the published NME2 cDNA sequence (Stahl *et al*, 1991), oligonucleotide primers (A and B) were designed from 5' and 3' coding regions of NME2 which demonstrated reduced nucleotide homology with the corresponding region in NME1 (described in section 2.6.2.1). Using this primer pair, the predicted 507bp product of the coding sequence was PCR amplified from a human HeLa cDNA library (as described in section 2.6.2.1: data not shown). This product was subcloned by TA

vector cloning (as described in section 2.9) and sequenced (as described in section 2.10). Sequencing revealed 100% nucleotide identity with that of the published sequence of NME2 (data not shown). The nucleotide homology of this region between NME1 and NME2 was found to be 74% (data not shown). This 507 bp DNA fragment was used as a gene specific probe for chromosomal localisation of NME2 (section 3.9).

3.9. Localisation of NME2 by Southern hybridisation

The 507 bp NME2 fragment was used to probe a southern blot of EcoRI digested DNA from a variety of human-rodent somatic cell hybrids (Black *et al*, 1993) and human and mouse controls (as described in section 2.5). An example of such a southern blot is shown in Figure 3.8(a). A complex pattern of fragments was detected in the human and mouse DNA controls (Figure 3.8(a) lanes 10 and 9 respectively). A 24 kb fragment in the human DNA control was distinguishable from a number of cross-hybridising mouse fragments. This fragment was present in the hybrid PCTBA1.8 (Figure 3.8(a) lane 1) which contains chromosome 17 as its only human contribution. A series of hybrids containing fragments of chromosome 17 were also analysed and the results are shown in lanes 2-8 (Figure 3.8(a)). The presence of the human-specific 24kb fragment in a proportion of these hybrids suggested the localisation of NME2 to the region 17q21-q22, the same localisation as NME1. The localisation is summarised in an ideogram of chromosome 17 in Figure 3.8(b).

However due to the ambiguity of the fragment pattern in the human and murine DNA, in particular, the presence of a 13kb human-specific fragment (tracks 4 and 5, Figure 3.8(a)) which does not map to chromosome 17, and the previous suggestion that an NME2 gene may be located on chromosome 16 (Steeg and Liotta, 1990), a second mapping strategy was employed to determine unequivocally that NME2 maps to 17q21-q22 (see following section).

Figure 3.8. Southern hybridisation of NME2 to chromosome 17q21-22.

The 507 bp NME2 fragment (described in section 3.8) was used to probe a southern blot of EcoRI digested DNA from a variety of human-rodent somatic cell hybrids.

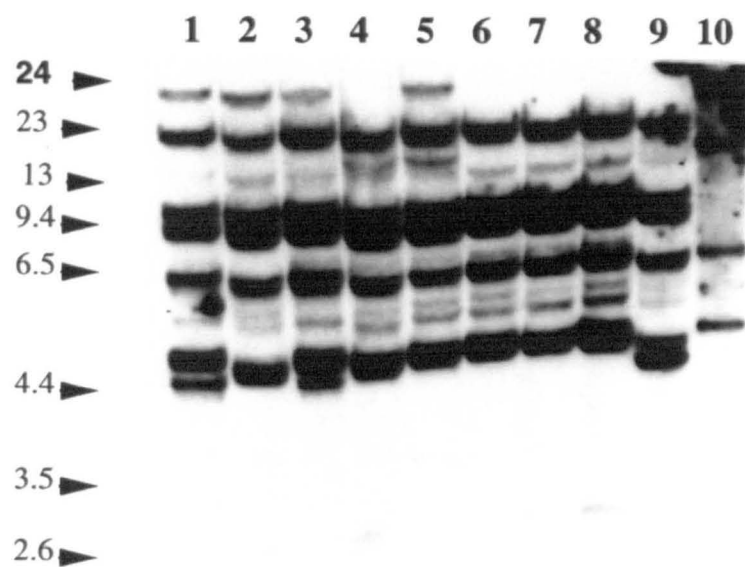
(a) An autoradiograph of somatic cell hybrid and control DNA samples hybridised with the NME2 fragment.

Marker sizes (in kb) indicated on the left-hand side are estimated from λ DNA digested with Hind III. The approximately 24 kb human specific fragment is also indicated on the left of the autoradiograph. The lane order and approximate human chromosome 17 content (Black *et al*, 1993) is: lane 1) PCTBA1.8 (17 only), lane 2) TRID62 (17q), lane 3) pJT2/A1 (17q12-qter), lane 4) p12.3B6 (pter-17q12), lane 5) DCRI (17q11-qter), lane 6) KLT8, lane 7) PLT6B, lane 8) PLT8 (KLT8, PLT6B and PLT8 are chromosome-mediated gene transfectants and contain multiple fragments of chromosome 17; however, PLT8 appears to have 17q23-qter as its only human chromosome 17 complement), lane 9) mouse and lane 10) human DNA. Lanes 1, 2,3,5 and 10 have the 24 kb human NME2-specific fragment.

(b) Ideogram of chromosome 17 showing localisation of NME2 to 17q21-22.

The approximate positions of the chromosome 17 bands are shown to the left of the chromosome ideogram. To the right of the ideogram, the approximate chromosome 17 content of four of the hybrids is indicated. (a) PCTBA1.8, (b) TRID62, (c) PLT8, (d) PLT6B. The localisation of NME2 on chromosome 17 is shown to the right of the hybrids. The human NME2-specific 24 kb fragment was detected in the hybrids (a) and (b) but not in (c) and (d); localising NME2 to 17q21-22.

(a)



(b)

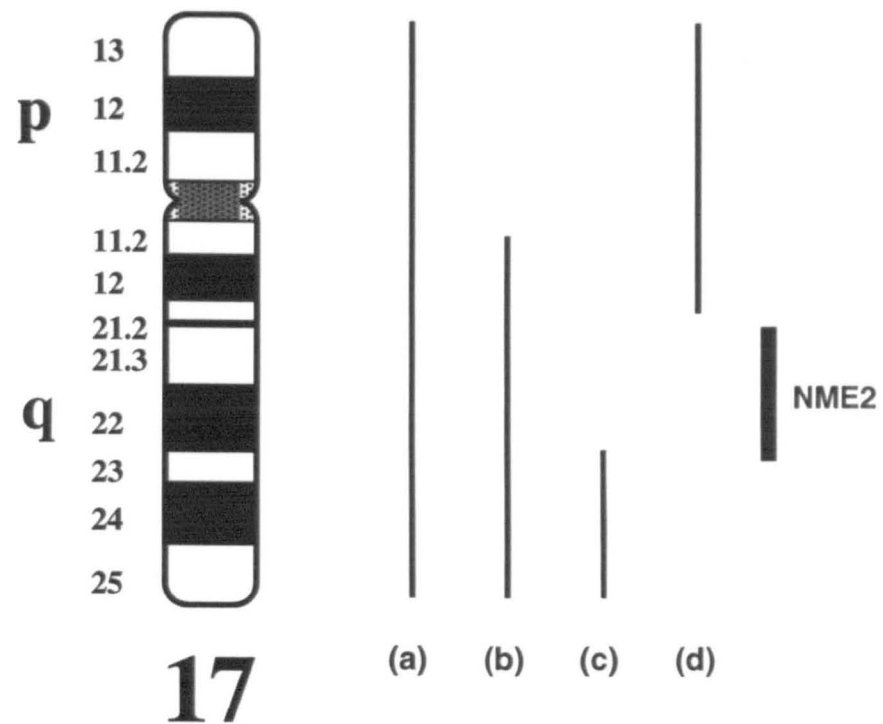


Figure 3.8. Southern hybridisation of NME2 to chromosome 17q21-22

3.10. Localisation of NME2 to 17q21-22 by PCR amplification

From the published DNA sequence (Stahl *et al* 1991), oligonucleotide primers were designed from the 3' untranslated region of the human NME2 gene. This region was chosen on the basis of the following observations:

- (i) comparison of the cDNA sequences of this region of NME1 and NME2 identified a number of sequence mismatches allowing NME2-specific oligonucleotide primers to be designed
- (ii) the 3' untranslated regions of genes are often under reduced interspecies sequence conservation suggesting NME2 oligonucleotide primers designed from this region would be less likely to PCR amplify rodent NME2 (the DNA background of the somatic cell hybrids is rodent)

The position of the NME2 oligonucleotide primers, indicated as C and D, used in the human/rodent somatic cell hybrid mapping are shown in Figure 3.9(a). The sequence comparison of NME2 primer sequences with the equivalent region in NME1 and NME2 is also shown. PCR amplification from a panel of monochromosomal human/rodent somatic cell hybrids (Kelsell *et al*, 1995a) and hybrids containing a number of human chromosomes (Spurr *et al*, 1991) was performed using the human NME2 specific oligonucleotide primers C and D (as described in section 2.6.2.2). A NME2-specific PCR amplified product was obtained using human (lane 6) and a selection of somatic cell hybrid DNAs (lanes 2, 4, and 5) as templates (Figure 3.9(b)). Lane 2 contains DNA from the hybrid PCTBA1.8 which has chromosome 17 as its only human complement. Lanes 4 and 5 contain DNA from the hybrids Dur4.3 and 3W4CL5 respectively, which have chromosome 17 and other chromosomes as their human complements. To further localise NME2, the series of hybrids containing fragments of chromosome 17 used in the localisation of NME2 by southern hybridisation (previous section) were analysed by PCR amplification with primers C and D. The expected 169bp product was present in the following hybrids: PCTBA1.8, TRID62, pJT2/A1, DCRI (data not shown). These results confirm the localisation of NME2 to 17q21-q22 (Kelsell *et al*, 1993b).

Figure 3.9. Localisation of NME2 to 17q21-22 by PCR amplification

(a) NME2 oligonucleotide primers, C and D, used for PCR amplification.

The position of each primer relates to its nucleotide position for the published sequence of NME2 (Stahl *et al*, 1991) and is shown above each primer sequence. The primer sequence is shown in bold: lower case indicates a different nucleotide in that position compared to NME1 (Steeg *et al*, 1988); uppercase indicates identical nucleotide to that of NME1.

(b) Mapping of NME2 to chromosome 17 by PCR amplification

Photograph of an ethidium bromide-stained 2% agarose gel showing the NME2-specific PCR amplified product. The 169 bp NME2-specific amplification product is indicated by an arrow to the left of the photograph and was present in lanes 2, 4, 5 and 6. The lane order is as follows: lane M) ϕ x174 DNA digested with *Hae*III ; lane 1) FG10; lane 2) PCTBA1.8, lane 3) 2860H7, lane 4) Dur4.3, lane 5) 3W4CL5, lane 6) human, lane 7) mouse, lane 8) hamster. The human chromosome complement of the somatic cell hybrids: FG10, Dur4.3 and 3W4CL5 are referenced in Spurr *et al*, (1991). Hybrid PCTBA1.8 contains chromosome 17 as its only human complement and 2860H7 contains only human chromosome 16 (Kelsell *et al*, 1995a).

Figure 3.9. Localisation of NME2 to 17q21-22 by PCR amplification

(a) **Primer C** (5' to 3' from position 500 to 522)

GtctTGTGCTCAtgACTGGgTC

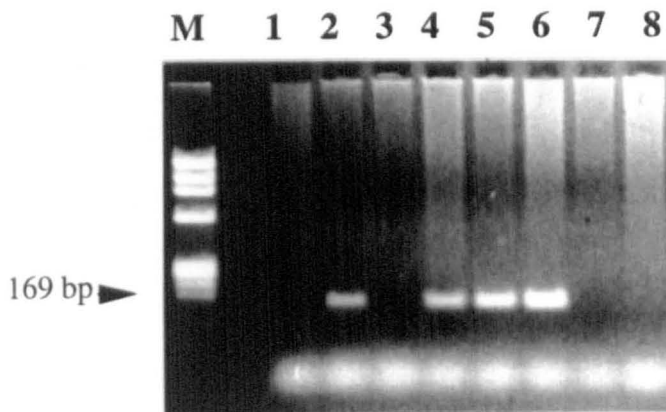
Mismatches with human NME1: **6/22** nucleotides

Primer D (3' to 5' from position 669 to 646)

cggctTCCAAaAGCTTtatTggc

Mismatches with human NME1: **12/23** nucleotides

(b)



However, from the genotyping of affected individuals from BRCA1-linked breast cancer families using a microsatellite marker locus mapping close to NME1 and NME2, recombination events between the disease and this marker locus have been identified (see Figure 3.10). These observations have genetically excluded NME1 and NME2 as candidates for BRCA1.

3.11. Genetic exclusion of other BRCA1 candidate genes

With the narrowing of the minimal region containing BRCA1 and the refinement of the genetic map, a number of candidate genes for BRCA1 have been eliminated. Analysis of recombinants in clearly linked extended pedigrees placed BRCA1 in the interval telomeric to THRA1 (Bowcock *et al*, 1993), D17S800 (Section 3.5 (b); Kelsell *et al*, 1993a) and D17S776 (Goldgar *et al*, 1994), and centromeric to D17S183 (Bowcock *et al*, 1993) and D17S78 (Simard *et al*, 1993). This region genetically excludes a number of candidate genes including NME1 and 2, ErbB2 which is amplified in a proportion of breast tumours, the putative oncogene THRA1 and Prohibitin (PHB), in which somatic mutations in a small proportion of sporadic breast tumours have been detected (King, 1992; Sato *et al*, 1992; Bowcock *et al*, 1993; Kelsell *et al*, 1993b)). The mapping positions of these and other excluded candidate genes are shown in Figure 3.10.

3.12. DNA mutation analysis of EDH17B1 and EDH17B2

Two 17 β -oestradiol dehydrogenase genes, EDH17B1 and EDH17B2, are both localised on a 13kb genomic DNA fragment (Luu-The *et al*, 1990) which maps within a defined minimal region for BRCA1 between D17S776 and D17S78 (Figure 3.10). 17 β -oestradiol dehydrogenase catalyses the interconversion between the weak oestrogen, oestrone, and the more potent oestradiol, the levels of which have been shown to be much higher in some breast tumours than in normal breast cells (McNeil *et al*, 1986; Vermeulin *et al*, 1986). Two mRNA species have been identified, one of 1.3 kb and another of 2.2 kb, in varying abundance in most tissues, including the breast and ovary (Luu-The *et al*, 1990). Nucleotide sequence comparison of the two

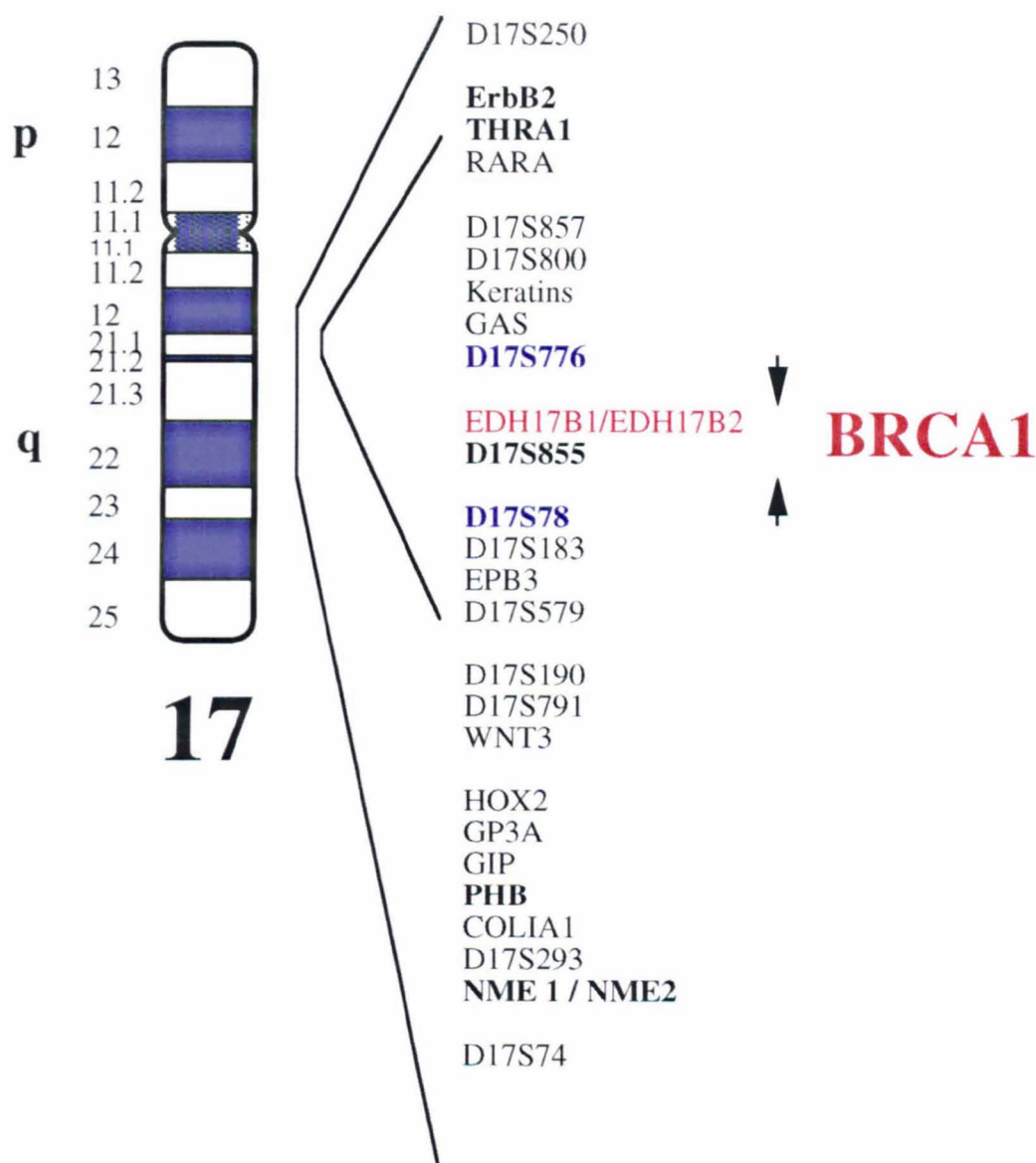


Figure 3.10. Exclusion of candidate genes for BRCA1

The order of loci mapping in the 17q12-22 interval are shown to the right of the ideogram of chromosome 17. D17S74, the DNA marker locus first indicating linkage with familial breast cancer, maps approximately 24 cM telomeric to D17S579 (Cox *et al*, 1994). The current minimal genetic region harbouring BRCA1 is between D17S776 and D17S78 (loci highlighted in blue). Genes previously implicated as candidate genes for BRCA1 including NME2 are in bold text and, as they map outside the minimal region, can be genetically excluded as candidates for BRCA1. The 17 β -oestradiol dehydrogenase genes (highlighted in red text), EDH17B1 and EDH17B2, are localised within the minimal region. The genetic location and biological role of these genes implicate them as plausible candidates for BRCA1.

genes over their exons and introns shows 89% homology. EDH17B1 has a stop codon at position 218 instead of the amino acid Gln as in EDH17B2 and is likely to be a pseudogene. However, it could still transcribe and encode a protein of 214 amino acids (Luu-The *et al*, 1990). The regulation of 17 β -EDH expression may be important in the initiation and/or progression of breast (and ovarian) tumours. The genetic location and biological role of these genes implicated them as potential candidates for BRCA1.

Sequence analysis of these candidate genes was performed on genomic DNA extracted from lymphoblastic cell lines from three affected individuals (3550, 3543 and 3677) and one unaffected non-gene carrier (3079) from the BOV3 pedigree. The whole coding region of both EDH17B1 and EDH17B2 was sequenced in an attempt to identify any disease-causing mutations in their constitutive DNA.

Oligonucleotide primers 5' and 3' to the coding region and specific for either EDH17B1 and EDH17B2 were designed from the published sequences (Luu-The *et al*, 1990). These are shown in Figure 3.11(a) and (b). The two gene specific templates produced using these primer sets following PCR amplification were added to a series of nested PCR amplification reactions (as described in section 2.6.3). The combination of internal primers (E to M) were used to amplify shorter PCR products which are more amenable for DNA sequencing. These primers are shown in Figure 3.11(b). Given the high degree of nucleotide homology between the two genes, these internal primers were designed from regions of high conservation so that identical sets of primers could be used to amplify each of the six exons from both genes. One of the internal primers from each pair was biotinylated to allow the isolation of single-stranded DNA and direct sequencing (described in section 2.10). An example of the specificity of the technique is shown in Figure 3.11(c), which shows partial sequence of exon 5 for both genes.

Sequencing of these genes revealed that there was no apparent variation between affected individuals and the unaffected individual from the family: no individual gave a sequence which differed from that of the published sequence. All individuals were homozygous for the published exon polymorphisms (Normand *et al*, 1993). In support of this data, EDH17B2 has been sequenced in a number of other BRCA1 gene carriers and no disease causing mutations have been identified (Simard *et al*, 1993). Complete confidence in the elimination of these genes as candidates for BRCA1 requires identifying a recombination event between this locus and BRCA1 or, indeed, identifying BRCA1.

3.13. Summary of Chapter 3

Linkage studies were performed on 16 breast and breast-ovary cancer families to confirm linkage to 17q and to identify closer flanking DNA marker loci to the disease susceptibility gene, BRCA1. The data was analysed independently and also in combination with other family data from the Breast Cancer Linkage Consortium group. In the largest of the ICRF families studied (BOV3), linkage to the long arm of chromosome 17 was confirmed and a number of recombinants were identified. One such cross-over event (identified in individual 3684) enabled the reduction of the interval harbouring BRCA1 to a region estimated to be between 1-1.5 Mb. Using this information a number of candidate genes for BRCA1 could be genetically excluded. Sequence analysis did not identify any disease-causing mutations in the genes, EDH17B1 (which codes for 17 β -oestradiol dehydrogenase) and EDH17B2 which are localised within the published minimal genetic region for BRCA1.

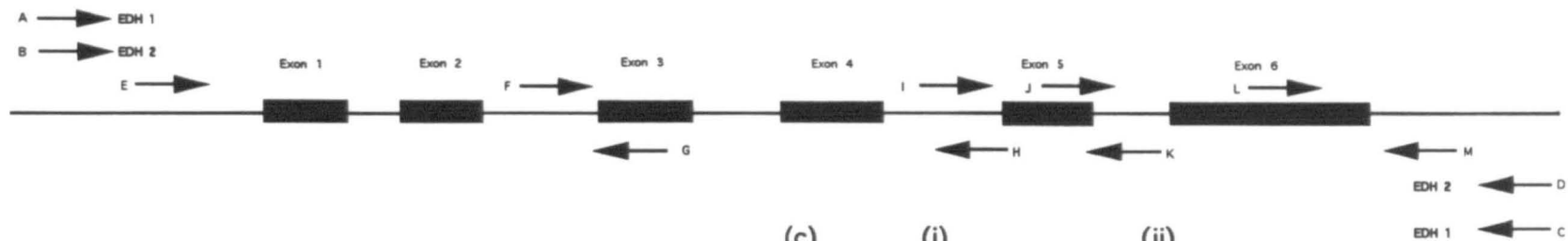
Figure 3.11. Sequencing of EDH17B1 and EDH17B2.

(a) A schematic representation (not to scale) of the structure of the EDH17B1 and EDH17B2 genes is shown. The exons are boxed and are labelled above each box. The positions and orientation of the oligonucleotide primers used in the gene-specific PCR amplification and sequencing are indicated by labelled arrows.

(b) Primer sequences used in the PCR amplification and sequencing of the two genes are shown. The gene-specific primer sequences are indicated (A, B, C and D) which were used to PCR amplify gene-specific template. The asterisk (*) indicates that the position of the primer sequences are relative to the first nucleotide of the translation initiation codon labelled 1.

(c) Examples of autoradiographs from sequence analysis of the same region of exon 5 for each gene: (i) EDH17B1 and (ii) EDH17B2. The specific dideoxy terminator is indicated above each sequencing lane.

(a)



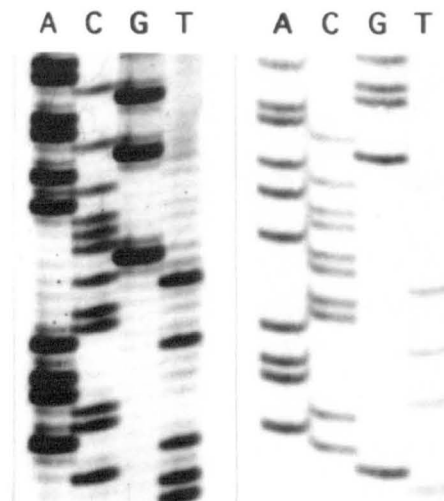
(b)

Primer	Sequence	Nucleotide position*
A (EDH1)	5' GGATGAGTTCTTCACCTTCTG	-810 -789
B (EDH2)	5' CTCCTGGGCTCCAACAATCC	-803 -783
C (EDH1)	5' GCCGCGCAGATATAACGGGC	2137 2117
D (EDH2)	5' GCTGCGCTGGTAAACTGGC	2138 2118
E	5' GCGAAGCAGGTGATATCAAGC	-54 -33
F	5' GCTGAGGGTGATGCTGAGGC	458 478
G	5' CAGGCCCAGGCCTGCGTTAC	527 507
H	5' TGCTCGGCATCTCTAGGTCAG	1011 990
I	5' GTTCTGGTTATCCCCAGGC	1310 1330
J	5' ACCGCCTTCATGGAGAAGGT	1501 1521
K	5' CACCTCCTCAGGGTTCTGCG	1643 1623
L	5' ACTACGTCACCGCCATGCAC	1841 1861
M	5' CAGACCCAGGGCACAAGAA	2055 2035

(c)

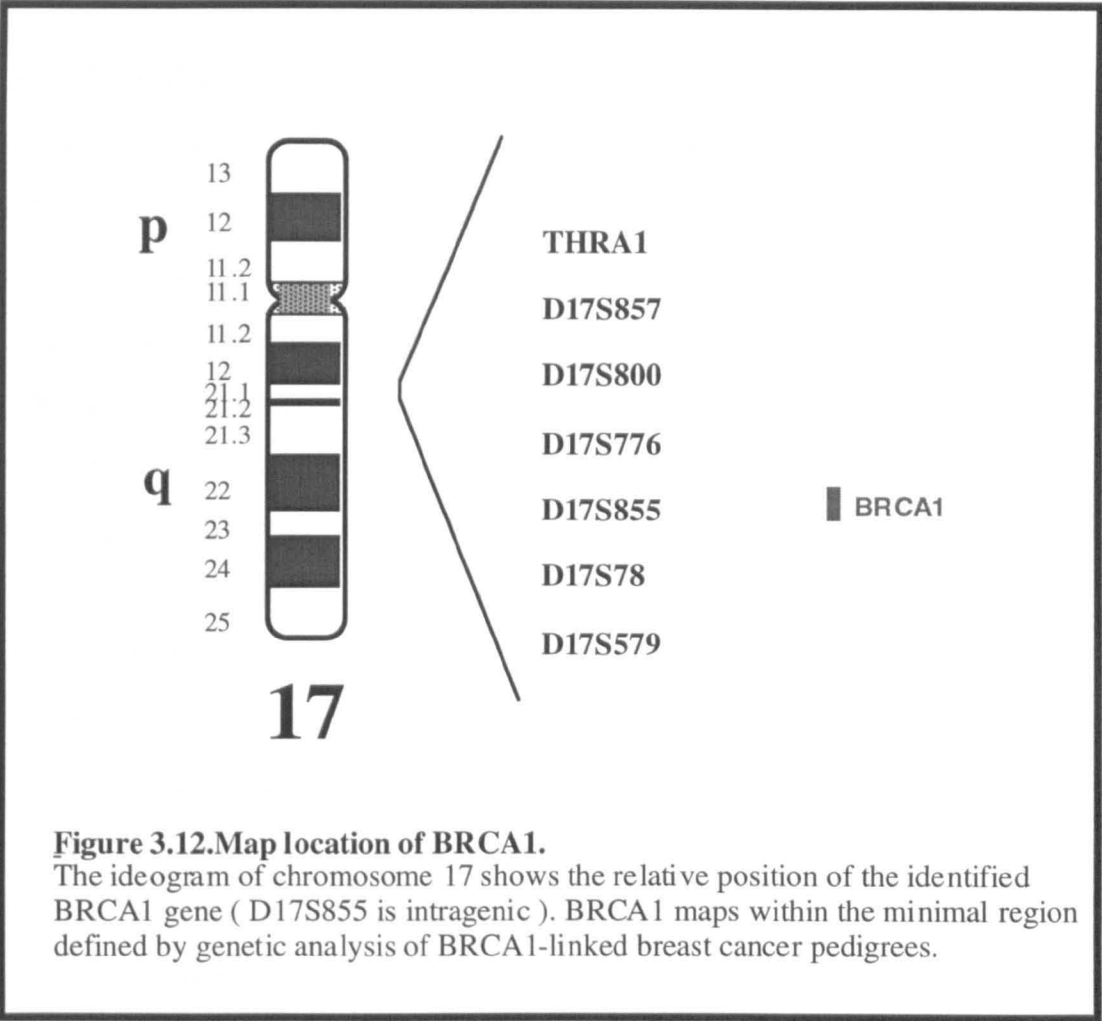
(i)

(ii)



3.14. The cloning of the BRCA1 gene

After an intense search in this gene rich region by a number of research groups worldwide, BRCA1 was identified (Miki *et al*, 1994). The gene consists of 22 coding and two non- coding exons spanning a genomic region of around 100 kb and encompasses the micosatellite marker locus, D17S855 (Figure 3.12).



The resulting 1863 amino acid protein bears little homology to any other known protein. However in the amino terminus of the protein, there is a putative RING finger (zinc- binding) domain which suggests a transcriptional regulatory role for BRCA1. Thompson *et al* (1995) proposed that BRCA1 may act as a growth suppressor of breast epithelial cells, though other functional studies are required. Previous studies had indicated that BRCA1 acts as a tumour suppressor gene (see Chapter 5). Developmental studies of murine BRCA1 in the mammary gland suggest its

expression is induced in response to ovarian hormones at, for example, puberty, and is involved in proliferation and differentiation in breast epithelial cells (Marquis *et al*, 1995). As BRCA1 is expressed in a number of tissues other than the breast and ovary, studies are required to determine why BRCA1 mutations confer cancer susceptibility in such a tissue specific manner. The possible function of BRCA1 is discussed further in Chapter 7.

Over 100 BRCA1 mutations have been identified which cover a wide spectrum of the gene (Miki *et al*, 1994; Shattuck- Eidens *et al*, 1995; Gayther *et al*, 1995; Friend *et al*, 1995). Over 75 % of detected mutations result in a truncated BRCA1 protein. Several common mutations have been identified, in particular, a 2 bp deletion resulting in a frame shift at position 185 in exon 2 (185delAG). This mutation appears to have arisen independently in a number of populations but a proposed ancestral mutation has a carrier frequency of around 1 % in the Ashkenazi Jewish population (Streuwing *et al*, 1995a). Interestingly, the same mutation is the genetic basis of disease predisposition in the BOV3 family but, from haplotype analysis, arose independently to that found in the Ashkenazi Jewish population (Neuhausen *et al*, 1996a). Other BRCA1 mutations detected in the ICRF pedigrees are discussed in Chapter 5. There is the suggestion that there may be a genotype- phenotype correlation; with the risk of ovarian cancer being reduced if the mutation occurs in the 3' third of BRCA1 (Shattuck-Eidens *et al*, 1995; Gayther *et al*, 1995). Two population based surveys to estimate the prevalence of BRCA1 mutations in women diagnosed with breast cancer at an early age have been performed (Fitzgerald *et al*, 1996; Langston *et al*, 1996). Minimum estimates suggest that between 10-15 % of early onset breast cancer cases carry a germline mutation in BRCA1.

CHAPTER 4: PHYSICAL MAP OF THE BRCA2 REGION

4.1. Introduction

Linkage studies have demonstrated that the majority of families with multiple cases of breast and ovarian cancer, but only around 45% of breast cancer only families, are due to inherited mutations in BRCA1 (Easton *et al*, 1993b; Narod *et al*, 1995). Examples of ICRF families in which the cancer is likely to be unlinked to BRCA1, from haplotype and multi-point linkage analysis (as described in section 2.13.1.2) with microsatellite marker loci flanking the BRCA1 locus, are shown in Figures 4.1 and 4.2. These observations suggested that there is at least one other major locus for breast (and ovarian) cancer predisposition. Steichen-Gersdorf *et al* (1994) demonstrated evidence in favour of linkage of the disease to BRCA1 in seven site specific ovarian cancer families; however, OV579 is a site-specific ovarian cancer family (all three cases diagnosed under the age of 40 years) in which ovarian cancer does not segregate with a BRCA1 haplotype (Figure 4.1(i)).

From a genomic linkage search using 15 families unlinked to the BRCA1 locus, a second breast cancer susceptibility locus, BRCA2, was localised to a 6 cM interval on 13q12-13, flanked by D13S289 and D13S267 (Wooster *et al*, 1994). Interestingly, a previous study excluded linkage of familial breast cancer to the retinoblastoma gene locus (RB1) (Bowcock *et al*, 1990). RB maps approximately 20 cM distal to the most telomeric marker locus, D13S267, for BRCA2 (Kooy *et al*, 1994, Dib *et al*, 1996).

The penetrance of breast cancer among female BRCA2 carriers is similar to that for BRCA1 carriers; estimated to be about 87% from segregation analysis of a large breast cancer family in which the disease is linked to BRCA2 (Bishop *et al*, 1988; Wooster *et al*, 1994). However, BRCA2 appears to confer a lower risk of ovarian cancer, but a much higher risk of male breast cancer compared to BRCA1. Male breast cancer is a hundred times less common than that of female disease in the general population. (Germline mutations in the chromosome X-linked androgen receptor gene have been

Figure 4.1. Three breast and breast-ovarian cancer families not linked to BRCA1.

Three examples of ICRF breast and breast-ovarian pedigrees which appear unlinked to BRCA1 from haplotype and linkage analysis. The sex of each individual is as described in Figure 1.1. Filled boxes indicate affected individual; red=ovarian cancer; blue = breast cancer. Underneath each individual is their identification number plus either BC=breast cancer or OV=ovarian cancer (if affected) and the age of diagnosis of the cancer (if known). Haplotypes are shown underneath each individual for which DNA was available with microsatellite marker loci order (centromere to telomere) tested in each family to the right of each pedigree. LOD scores at the BRCA1 locus are shown under the pedigree identification number and were calculated as described in section 2.13.1.2.

(i) OV579. This is a site specific ovarian cancer family. A negative LOD score of -0.90 at BRCA1 suggests disease susceptibility in this family is not linked to BRCA1. This is clearly demonstrated from haplotype comparison of 15709 and 15697. Affected individual 15697 appears to have inherited the haplotype (1:3:3:1:1) from her father (as inferred by the observation that her uncle 15703 has the same haplotype). This haplotype has not been transmitted to her affected cousin 15709.

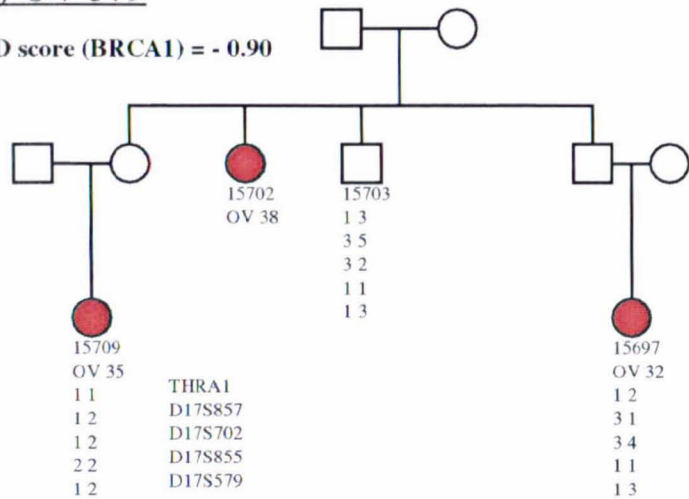
(ii) BC3. This is a breast-ovarian cancer family with one case (111) of bilateral breast cancer. A negative LOD score of -0.27 at BRCA1 suggests disease susceptibility in this family is not linked to BRCA1. This again can be demonstrated from comparison of haplotypes between affected members (120, 116, 111 and 145) and the obligate gene carrier (140). A haplotype of 1:2:1 is shared by 140, 116, 111 but not 120 or 145 or, alternatively, a haplotype of 2:1:1 is shared by 111, 116 and 120 but not 140 and, therefore, not 145.

(iii) BC12. This is a site specific breast cancer family consisting of four affected individuals. A negative LOD score of -0.58 at BRCA1 suggests disease susceptibility in this family is not linked to BRCA1. This is reflected by the observation that 3349 does not share a BRCA1 haplotype with her affected cousins, 3352 and 3354.

Figure 4.1. Three breast and breast-ovarian cancer families not linked to BRCA1.

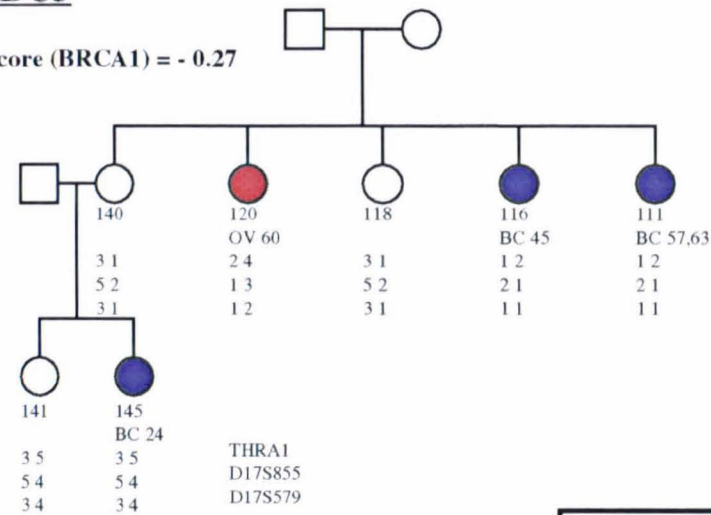
(i) OV 579

LOD score (BRCA1) = - 0.90



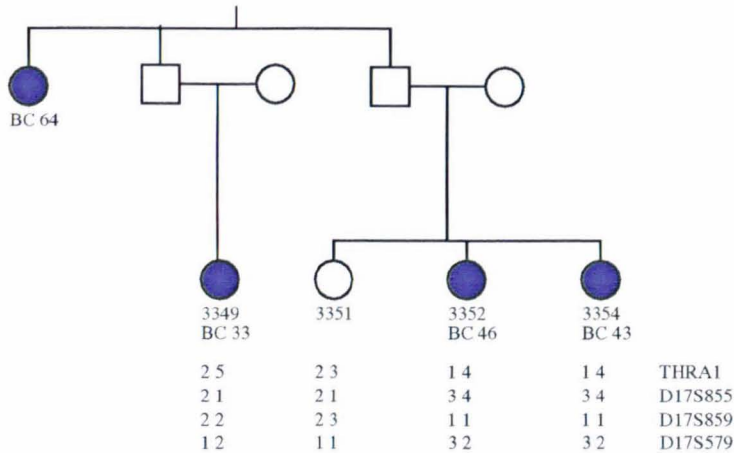
(ii) BC3

LOD score (BRCA1) = - 0.27



(iii) BC12

LOD score (BRCA1) = - 0.58



Key:

- ovarian cancer
- breast cancer

Figure 4.2. ICRF 26, a breast cancer family not linked to BRCA1.

The pedigree is constructed as described in Figure 4.1. Cancer at other sites are present in this family indicated by green filled box (circle for females, square for males); the specific cancer is described under the individual. A zero allele indicates lack of PCR product. Haplotypes for affected individuals: 17209, 17405 and 20576 were genotyped using DNA extracted from archival pathological tumour material (as described in section 2.3.3). A negative LOD score of -1.47 (calculated as described in section 2.13.1.2) suggests disease susceptibility in this family is not linked to BRCA1. There is no evidence of BRCA1 haplotype sharing amongst breast cancer affecteds in this family except for the sisters, 17199 and 17209 who share the haplotype: 1:1:4:5. However, this is of paternal origin and not from their affected mother 17187. For affected individual 18080, it is possible that a breast cancer susceptibility mutation may have been inherited paternally (father's sister had breast cancer diagnosed at 37 years) or maternally.

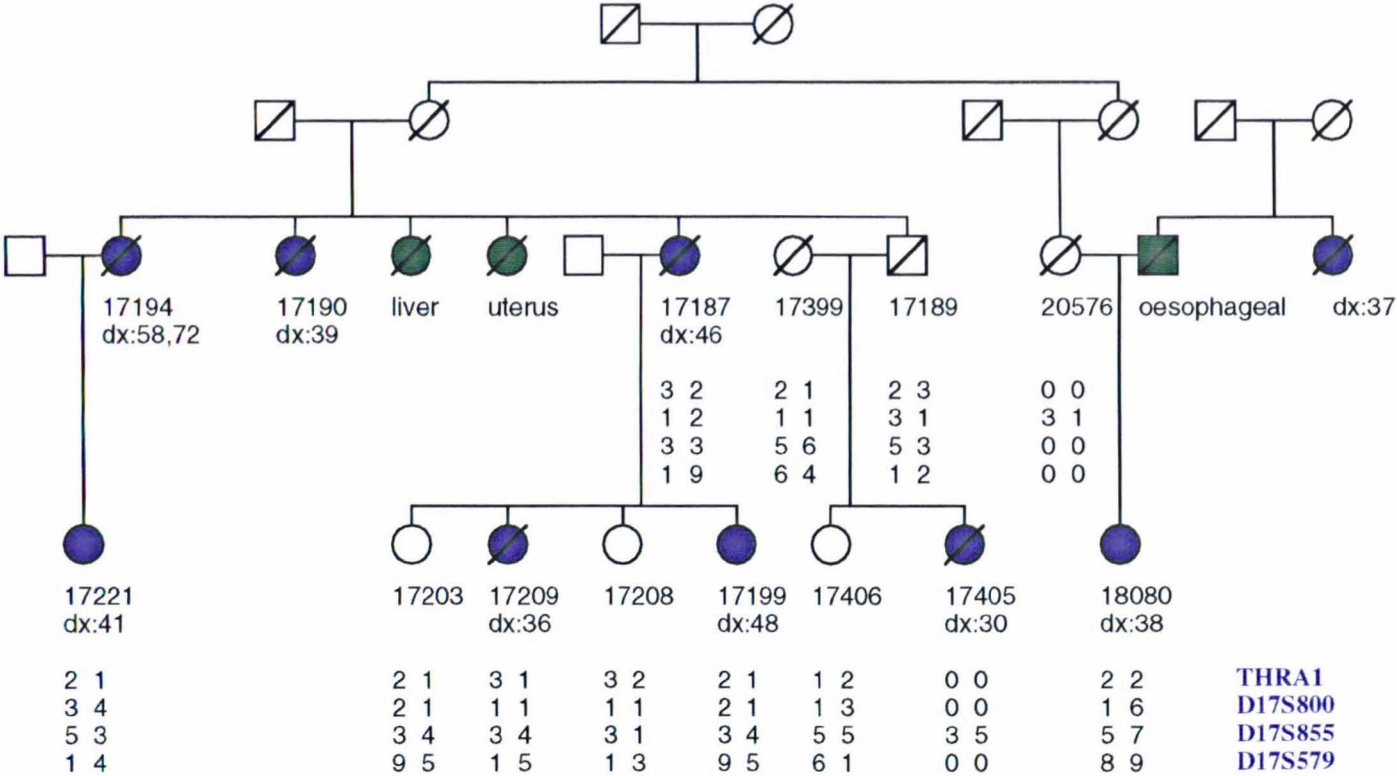
Figure 4.2. ICRF 26, a breast cancer family not linked to BRCA1.

LOD score (BRCA1) = -1.47

Key:

Breast cancer

Other cancer



demonstrated in males with breast cancer and androgen insufficiency: Wooster *et al*, 1992; Lobaccaro *et al*, 1993). Studies of families, in which male and female breast cancer was segregating, indicated that in the majority of these, the disease was unlinked to BRCA1 (Stratton *et al*, 1994). Subsequently, there have been reports of BRCA1 mutations segregating with breast cancer in a small number of families with a single case of male breast cancer (for example, Streuwing *et al*, 1995b). Other reports have confirmed linkage of breast cancer susceptibility to BRCA2 in a family with an excess of male breast cancer (Thorlacious *et al*, 1995) and a breast cancer family with cancer at other sites (Tonin *et al*, 1995).

In this study, a physical contig of the BRCA2 region was constructed based on a series of overlapping Yeast Artificial Chromosomes (YACs) containing human DNA. (The YAC (Yeast Artificial Chromosome) cloning system is capable of cloning large DNA fragments. The YAC vector contains the requirements for normal chromosomal replication including a centromere sequence, a telomere sequence, an autonomous replicating system and an origin of replication: reviewed in Strachan and Read, 1996.) The contig provided a resource for identifying new genetic markers and candidate genes for BRCA2.

4.2. Physical map of the BRCA2 region

Fourteen YAC clones containing fragments of human DNA tentatively mapping in the BRCA2 interval were selected directly (as described in section 2.11.1). Human DNA inserts had been sized previously by pulsed-field gel electrophoresis (Cohen *et al*, 1993). DNA extracted from these YACs were used as hybridisation probes to normal human lymphocyte metaphase spreads by fluorescence *in situ* hybridisation (FISH) to confirm their localisation to 13q12-13 (S. Ford, personal communication). From this analysis, two clones, 851_b_2 and 793_b_6, were shown to be chimeric and were discarded from the study.

DNA extracted from the non-chimeric YACs (as described in section 2.3.4) were screened by PCR amplification with microsatellite marker loci mapping within or just outside the minimal genetic location of BRCA2 (as described in section 2.6.1). All YACs tested contained at least one of the microsatellite marker loci. A typical result of the mapping of microsatellite marker loci sequences to the YAC contig is shown in Figure 4.3(a). The results are summarised in Table 4.1.

Initially, two minimal tiling path contigs were generated on the basis of their microsatellite marker loci content, one linking D13S289 with D13S260, and the other linking D13S171 with D17S267 (Figure 4.3(b)). An STS, STS1002, isolated from YAC 746_g_10 (Dr. R. Wooster, personal communication) was screened across the YACs using PCR amplification (as described in section 2.6.1). STS1002 was found to be present in the YACs 931_f_4 and 746_g_10. This result physically links these two YACs and the microsatellite marker loci, D13S260 and D13S171, and is shown in a graphic representation of the contig in Figure 4.4. Adding the human DNA insert sizes of the smallest YACs (910_h_2, 931_f_5, 886_d_8, 951_a_3) spanning the 6 cM region between D13S289 and D13S267, the region can be no more than 5.27 Mb. The exact physical size of the interval cannot be calculated until the precise positions of the overlaps are determined.

Table 4.1. Summary of marker loci content of YACs in the BRCA2 region by PCR amplification

YAC	LOCUS	SIZE (in kb)
910_e_11	D13S289	1720
940_g_4	D13S289 D13S290 D13S1226	1170
962_a_8	D13S289 D13S290 D13S1226	1210
910_h_2	D13S289 D13S290 D13S1226	1030
931_f_4	D13S1226 D13S260 (STS1002)	1520
931_f_5	D13S1226 D13S260 (STS1002)	1510
833_d_2	D13S260	810
834_c_8	D13S260	750
851_b_2*	D13S260	400
793_b_6*	(STS1002) D13S171	400
746_g_10	(STS1002) D13S171	1650
886_d_8	(STS1002) D13S171	970
951_a_3	D13S171 D13S267	1760
941_d_4	D13S267	1580

STS1002 is shown in brackets as it was isolated from YAC 746_g_10 DNA after the initial microsatellite marker loci screen of the YAC DNA panel (Figure 4.4).

* YACs found to be chimeric

Figure 4.3. YAC characterisation of the BRCA2 minimal region

This figure shows the initial analysis of the 12 non-chimeric YAC DNAs selected, which tentatively mapped to the BRCA2 region. These were screened for presence or absence of specific microsatellite marker loci by PCR amplification.

(a) D13S260 PCR screen of the YAC DNA panel.

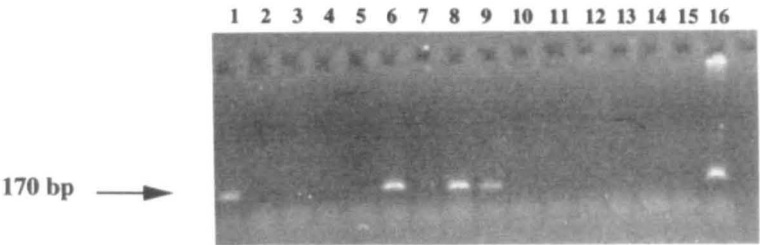
D13S260 PCR-specific products (sized previously as described in section 2.7.3) were obtained in lanes 1 (931_f_4), 6 (833_d_2), 8 (931_f_5), 9 (834_d_2) and 16 (human genomic DNA control). No PCR product was detected in the remainder of the lanes: 2 (941_d_4), 3 (910_e_11), 4 (951_a_3), 5 (886_d_8), 7 (910_h_2), 10 (962_a_8), 11 (746_g_10), 12 (940_g_4), 13 (total yeast/YAC genomic DNA). No DNA template was added to the PCR reactions loaded in lanes 14 and 15 (water controls).

(b) Graphic representation of results from microsatellite marker loci screen of YAC DNA panel

The order of STSs are shown positioned along the single line, including their orientation with respect to the centromere (CEN) and telomere (TEL) on the long arm of chromosome 13. The relative position and sizing of YACs (in brackets) is shown below this line. The estimated genetic distance between microsatellite marker loci in centimorgans (cM) is indicated above the map. Two contigs are indicated: one physically linking D13S289 with D13S260 and the other linking D13S171 with D13S267.

Figure 4.3. YAC characterisation of the BRCA2 minimal region

(a) D13S260 PCR screen of YAC DNA panel



(b) Graphic representation of results from microsatellite marker loci screen of YAC DNA panel

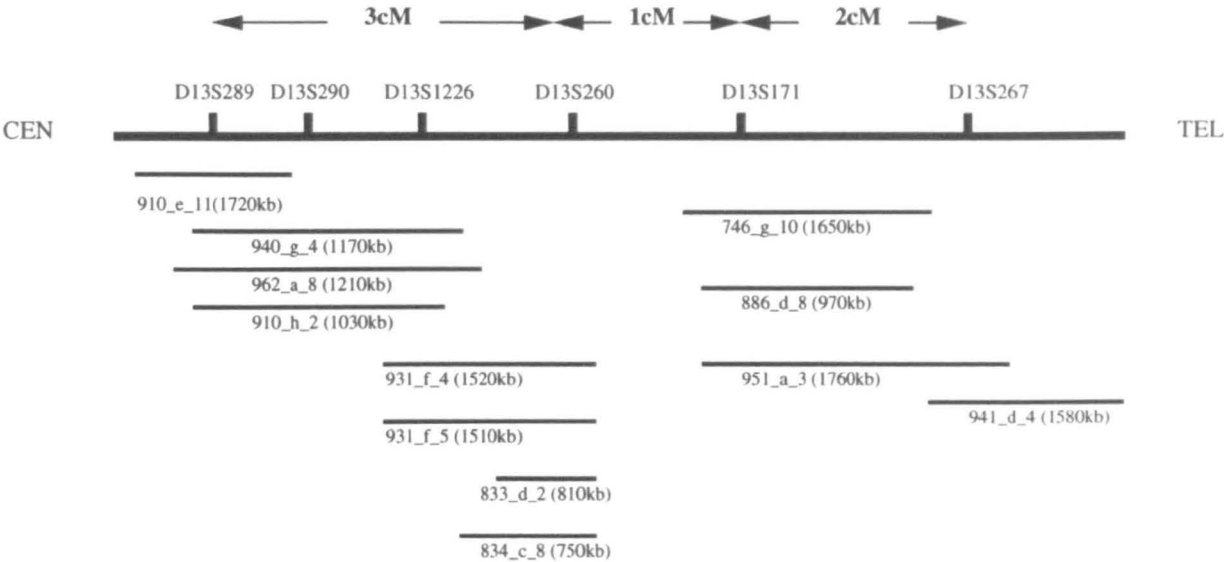
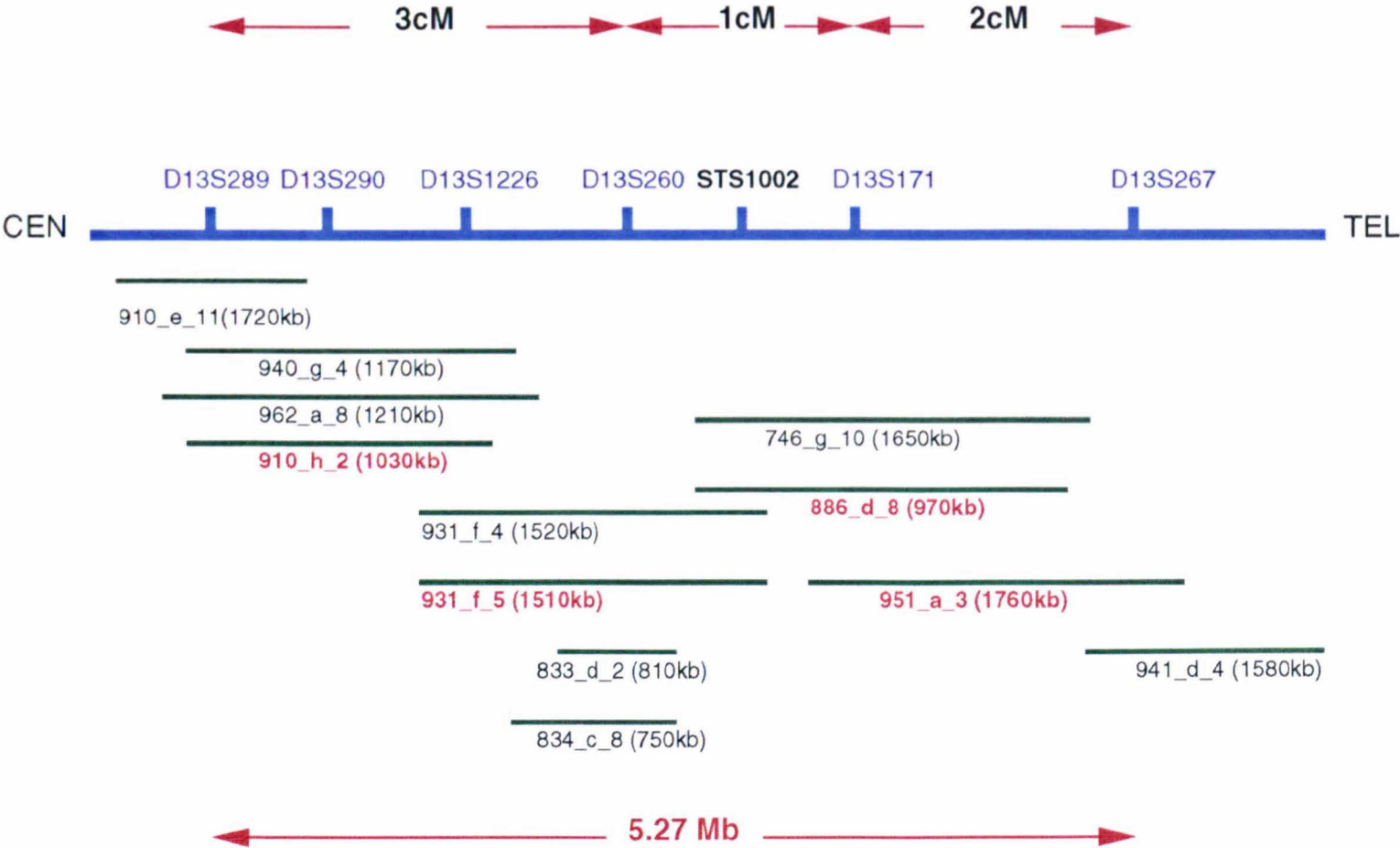


Figure 4.4. Graphic representation of the minimal tiling path of YACs spanning BRCA2 on 13q12-13.

The figure is as described in Figure 4.3(b), with the addition of colour. The order of marker loci are shown positioned along the blue line. The relative position and sizing of YACs is indicated by a green line. YACs highlighted in red are those used to estimate the approximate physical size of the region, 5.27 Mb (shown at bottom of the figure). A single contig spans the minimal published genetic interval for BRCA2, flanked by D13S289 and D13S267. The marker locus, STS1002 (highlighted in black above the blue line), links the two YAC contigs described in Figure 4.3(b).

Figure 4.4. Graphic representation of the minimal tiling path of YACs spanning BRCA2 on 13q12-13



4.3. Assignment of BRCA2 candidate genes

An analysis of the Genome Database (GDB; accessed via the HGMP computing facility) identified three possible candidate genes for BRCA2 mapping to the long arm of chromosome 13. These were the receptor-type tyrosine kinase genes, FLT1 and FLT3, and Brush1, a tyrosine kinase shown to be underexpressed in a proportion of breast tumours (Schott *et al*, 1994). FLT1-specific primers (Han *et al*, 1993) were used to determine if the FLT genes map to the contig; since it was known that FLT1 and FLT3 map within 600 kb of each other (Rosnet *et al*, 1993). Primers designed from the published sequence of Brush1 (as described in section 2.6.4) were also screened across the YACs. No positive PCR amplification signal for any of these three genes was present in the YACs, only in the human DNA control (data not shown). Recent physical and genetic maps from the chromosome 13 workshop confirm this result and position the two FLT genes and Brush1 centromeric to the BRCA2 gene region (Washington *et al*, 1995).

A number of partially sequenced, randomly isolated cDNAs have been identified (for example, Adams *et al*, 1995) and are referred to as expressed Sequence Tagged Sites (eSTSs). These sequences are usually from the 3' untranslated regions of genes, where there is less tendency for interspecies conservation and the insertion of intronic sequences. Two eSTS sequences, D13S178E and D13S181E, were shown to map to the chromosome interval on 13q which also harbours BRCA2 (Hawthorn and Cowell, 1996). Oligonucleotide primers designed for these transcribed gene sequences were screened across DNA from the YAC contig spanning the BRCA2 region by PCR amplification (as described in section 2.6.5). One of these genes, D13S178E, was found to map to the BRCA2 YAC contig. D13S178E-specific PCR products were obtained when the template DNA was isolated from the YACs 886_d_8 and 951_a_3, but not for the YACs 931_f_5 and 910_h_2 (Figure 4.5 (b)). Therefore, D13S178E maps telomeric to the microsatellite marker locus D13S260 (Figure 4.5.(a)). The position of D13S178E in relation to the microsatellite marker locus D13S171 could not be determined from present YAC data.

Figure 4.5. D13S178E, a candidate gene for BRCA2?

This figure shows the assignment of D13S178E to the BRCA2 region and a preliminary investigation by PCR amplification as to its pattern of expression.

(a) Assignment of D13S178E to the BRCA2 YAC contig.

This figure is as described in Figure 4.3 with the exception that only four YACs which span the region harbouring BRCA2 are shown below the single line. The relative position of D13S178E is shown above the line mapping to the region where D13S171 is located (the orientation of D13S178E to D13S171 cannot be determined from this analysis).

(b) Mapping of D13S178E to the BRCA2 YAC contig.

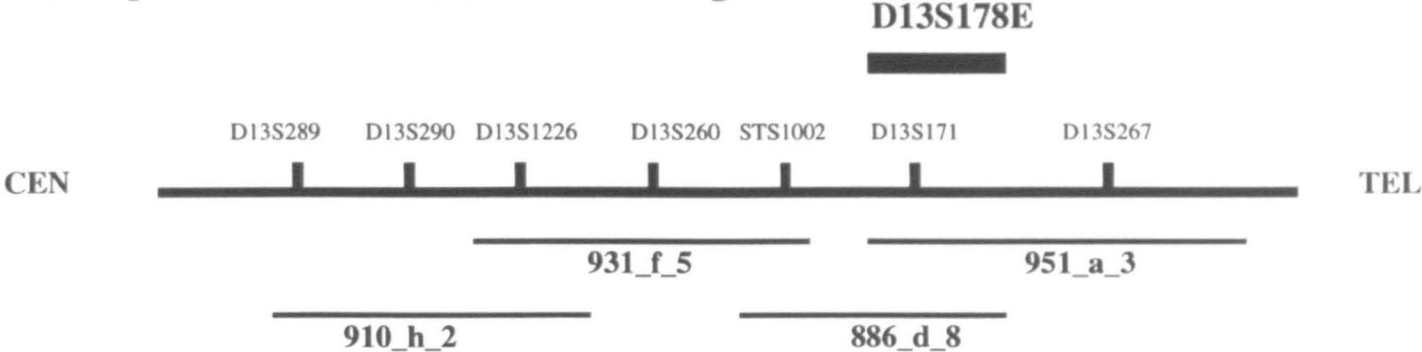
The expected 207 bp D13S178E-specific PCR product (when compared to the ϕ x174 DNA digested with *Hae*III molecular weight marker in lane M) was present in lanes: 1 (886_d_8 DNA), 2 (951_a_3 DNA) and 5 (human genomic DNA) but not detected in lanes: 3 (910_h_2 DNA) and 4 (931_f_5 DNA).

(c) D13S178E expression pattern.

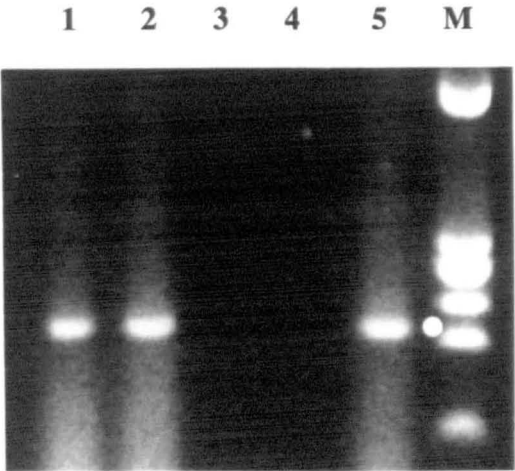
The 207 bp D13S178E specific PCR amplification product (size confirmed previously from Figure 4.5 (b)) were obtained in lanes: 1 (normal breast cDNA), 3 (foetal brain cDNA) and 7 (human genomic DNA). No PCR amplification product was detected in lanes: 2 (HeLa cDNA), 4 (placental cDNA), 5 (breast adenocarcinoma cDNA) and 6 (retina cDNA).

Figure 4.5. D13S178E, a candidate gene for BRCA2?

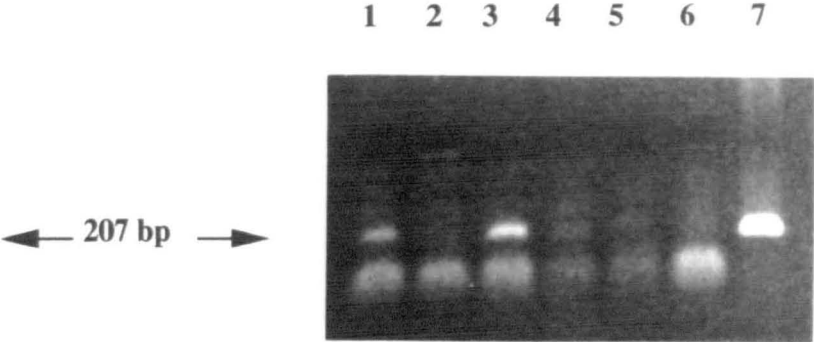
(a) Assignment of D13S178E to BRCA2 YAC contig



(b) Mapping of D13S178E to BRCA2 YAC contig



(c) D13S178E expression pattern



The sequence of D13S178E was screened across the GenEMBL sequence database (using the GCG package accessed via the ICRF computing facility) and revealed no homology to any previously identified genes or gene families (data not shown). To determine the expression pattern of D13S178E, the oligonucleotide mapping primers were screened across a panel of cDNA libraries by PCR amplification (as described in section 2.6.6). The 207 bp PCR product (confirmation of size determined from Figure 4.5 (b)) for D13S178E was present in normal breast and foetal brain cDNA libraries, but not in, for example, a breast adenocarcinoma cDNA library (Figure 4.5.(c)). The presence of a transcript in the normal breast but not in the breast adenocarcinoma library suggested that the full length cDNA of this gene could be regarded as a strong candidate for the proposed tumour suppressor gene, BRCA2. However, it is possible that the particular tissue-specific cDNA libraries used in this analysis were not fully representative of the number of genes expressed in those tissues.

4.4. Summary

A gene for breast/ovarian cancer predisposition, BRCA2, had been assigned to the region 13q12-13 (Wooster *et al*, 1994). Towards the identification of this gene, a YAC contig was constructed spanning the published minimal genetic interval between the marker loci D13S289 and D13S267. Physical mapping of these YACs using microsatellite marker loci suggested complete coverage of the region spanning a distance of 5.27 Mb or less. To identify candidate genes for BRCA2, the Genome Database was searched for genes mapping in this region. Three possible candidate loci: FLT1, FLT3 and Brush 1 were found to map outside the contig. One gene, D13S178E, was mapped to the contig and was regarded as a candidate for BRCA2.

This contig provided a framework for the generation of new polymorphic markers to assist in reducing the minimal genetic region for BRCA2, and for the isolation of other transcribed sequences. This work was continued to characterise the full length cDNA for D13S178E and to identify other candidate genes for BRCA2 (Drs. R. Wooster, M. Stratton and colleagues). Mutation analysis of DNA from BRCA2 gene carriers would

determine if D13S178E, or indeed another gene in the region, coded for the familial breast cancer gene, BRCA2.

4.5. Identification of BRCA2

The genetic interval harbouring BRCA2 was reduced to 600 kb from the identification of recombinants in BRCA2 linked breast cancer families (Wooster *et al*, 1995). One key observation came from the identification of a founder effect in a series of breast cancer families from Iceland (Gudmundsson *et al*, 1996). The region was reduced potentially to an interval of 300 kb just centromeric to D13S171, by the detection of a homozygous deletion in a pancreatic carcinoma (Schutte *et al*, 1995). The close proximity of this deletion with BRCA2 suggested the same gene may be involved in both cancer types.

A partial cDNA (not D13S178E) isolated from the minimal region for BRCA2 was identified in which six distinct germline mutations were found in BRCA2 linked breast cancer families (Wooster *et al*, 1995). As each mutation would disrupt the open reading frame and lead to a truncated protein, this cDNA was likely to be encoded by BRCA2. BRCA2 codes for a transcript of 10-12 kb which bears little homology to any known protein; though a region of the protein shows weak homology to a portion of the BRCA1 protein. The significance of this homology and other characteristics of BRCA2 require further elucidation (discussed further in Chapter 7).

(The families: OV579; BC3; BC12; ICRF26, described at the beginning of this chapter, show some evidence of linkage to BRCA2 but, as yet, no BRCA2 germline mutations have been detected (E. Mavraki, personal communication).

CHAPTER 5: ALLELE LOSS AT BRCA1 AND BRCA2

5.1. Introduction

This chapter describes allele loss studies with microsatellite marker loci flanking the BRCA1 and BRCA2 loci in breast carcinomas. These were performed on DNA extracted from tumours of putative BRCA1 gene carriers and a population of breast cancer patients unselected for family history. The majority of tumours from the latter group are assumed to have arisen without a highly penetrant inheritable susceptibility component, that is, they are sporadic. The initial aim was to investigate whether BRCA1 functions as a tumour suppressor gene in familial and sporadic disease. A number of these tumour DNA samples were also assessed for allele loss at the BRCA2 locus. Further genetic linkage studies incorporating tumour DNA analysis in ICRF breast cancer families are also described.

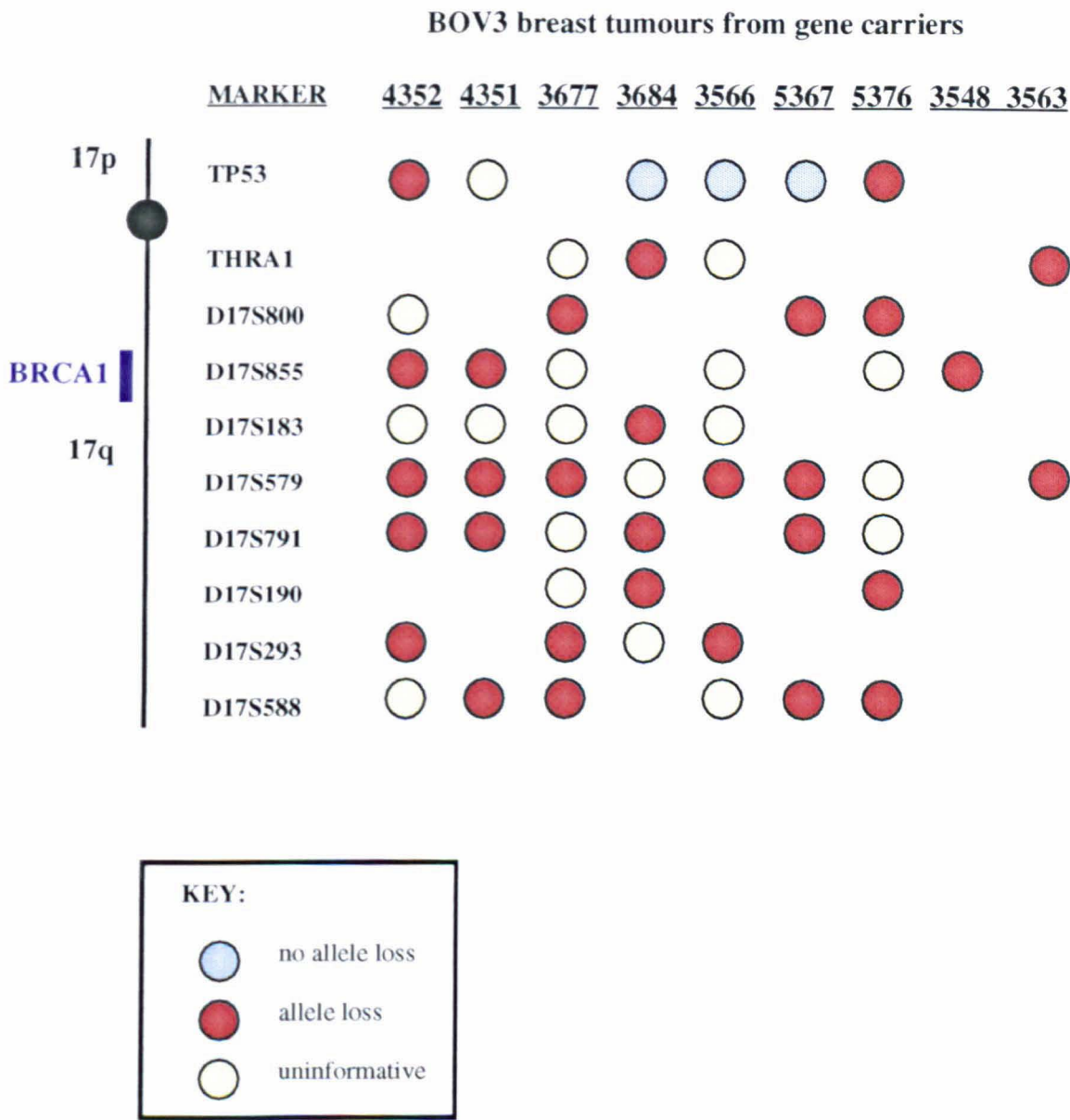
5.2. Is BRCA1 a tumour suppressor gene?

The focus for this study involved the analysis of carcinomas from the BOV3 pedigree (described in Chapter 3). In this family, an AG deletion in codon 185 of BRCA1 has been determined as the genetic basis of cancer predisposition (Shattuck-Eidens *et al*, 1995; Xu *et al*, in press). Allele status of BRCA1 in breast carcinomas from this family would provide convincing evidence as to a proposed tumour suppressor function for BRCA1. Breast tumour DNA was analysed using microsatellite marker loci flanking BRCA1. The nine breast carcinomas were from family members who were highly likely to be BRCA1 gene carriers from haplotype analysis of microsatellite marker loci flanking BRCA1 (Chapter 3 and subsequently confirmed from mutation analysis: E. Mavraki, personal communication). All tumour DNA analysed was extracted from archival pathological slide preparations (as described in section 2.3.3). The results for these nine carcinomas are shown in Figure 5.1; examples of allele status in tumour DNA are shown in Figure 5.2. Allele loss of the BRCA1 region was detected in all nine breast carcinomas with microsatellite marker loci localised to 17q12-21.

Figure 5.1. Allele loss on chromosome 17 in BOV3 breast tumours

This figure summarises the results obtained from the nine BOV3 breast carcinomas assessed for allele loss with microsatellite marker loci mapping to chromosome 17. All tumours shown in this figure are from BOV3 family members carrying the 185delAG mutation in BRCA1. The microsatellite marker loci screened across the tumours are shown to the left of the figure. The order of microsatellite marker loci is as described in Figure 3.10. All loci map to the region 17q12-21, with the exception of the TP53 microsatellite locus which maps on the short arm of chromosome 17. The pattern of allele loss/retention (as indicated from the key at the bottom of the page) is shown underneath the identification number of the individual (at top of figure) from which the tumour was derived. Lack of a symbol indicates allele status was not assessed at that locus. Each tumour displayed allele loss with the BRCA1 intragenic microsatellite marker locus D17S855, and/or microsatellite marker loci flanking BRCA1.

Figure 5.1 Allele loss on chromosome 17 in BOV3 breast tumours



(h)

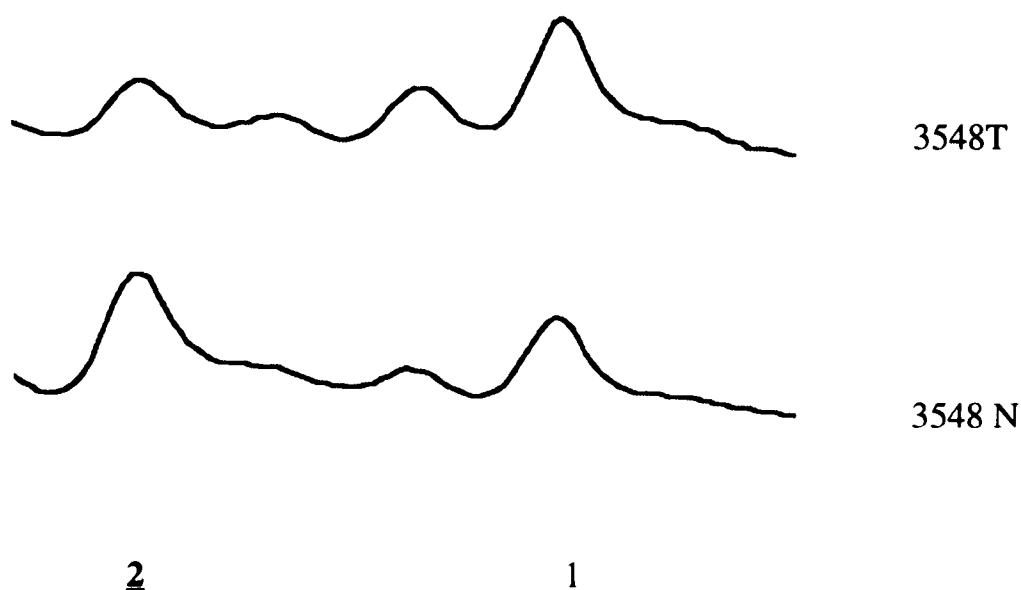


Figure 5.2. Examples of chromosome 17 allele status in breast carcinomas from BOV3

Examples of microsatellite allele loss analysis of tumour DNA from the BOV3 pedigree. The personal identification number is indicated above each lane (further individual details such as sex, disease and gene carrier status can be interpreted from Figure 3.4). T = tumour DNA, T1 = approximately 90 % tumour cell content (estimated as described in section 2.2.5), T2 = approximately 70 % tumour cell content, N = lymphocyte DNA. Allele typing is shown beneath each lane; allele in bold and underlined indicates the allele deleted in each tumour. For (h), the individual identification number and tumour type is shown to the right of each lane. Approximate allele size range (in base pairs: bp) in brackets after each marker locus.

(a) D17S293 (115-125 bp): allele 2 is from the disease haplotype.

(b) D17S800 (168-178 bp): allele 1 is from the disease haplotype.

(c) D17S293 (115-125 bp): allele 2 is from the disease haplotype.

(d) D17S855 (145-160 bp): allele 1 is from the disease haplotype.

(e) D17S190 (190-200 bp): allele 1 is from the disease haplotype.

(f) D17S183 (90-100 bp): allele 1 is from the disease haplotype.

(g) TP53 microsatellite locus in intron 4 (110-125 bp): representation of allele retention at this locus.

(h) D17S855 (145-160 bp): allele 1 is from the disease haplotype.

Allele detection in (a), (f) and (g) as described in section 2.7.1; in (b), (c), (d) and (e) as described in section 2.7.2; (h) as described in section 2.7.3. Allele loss was assessed as described in section 2.8.

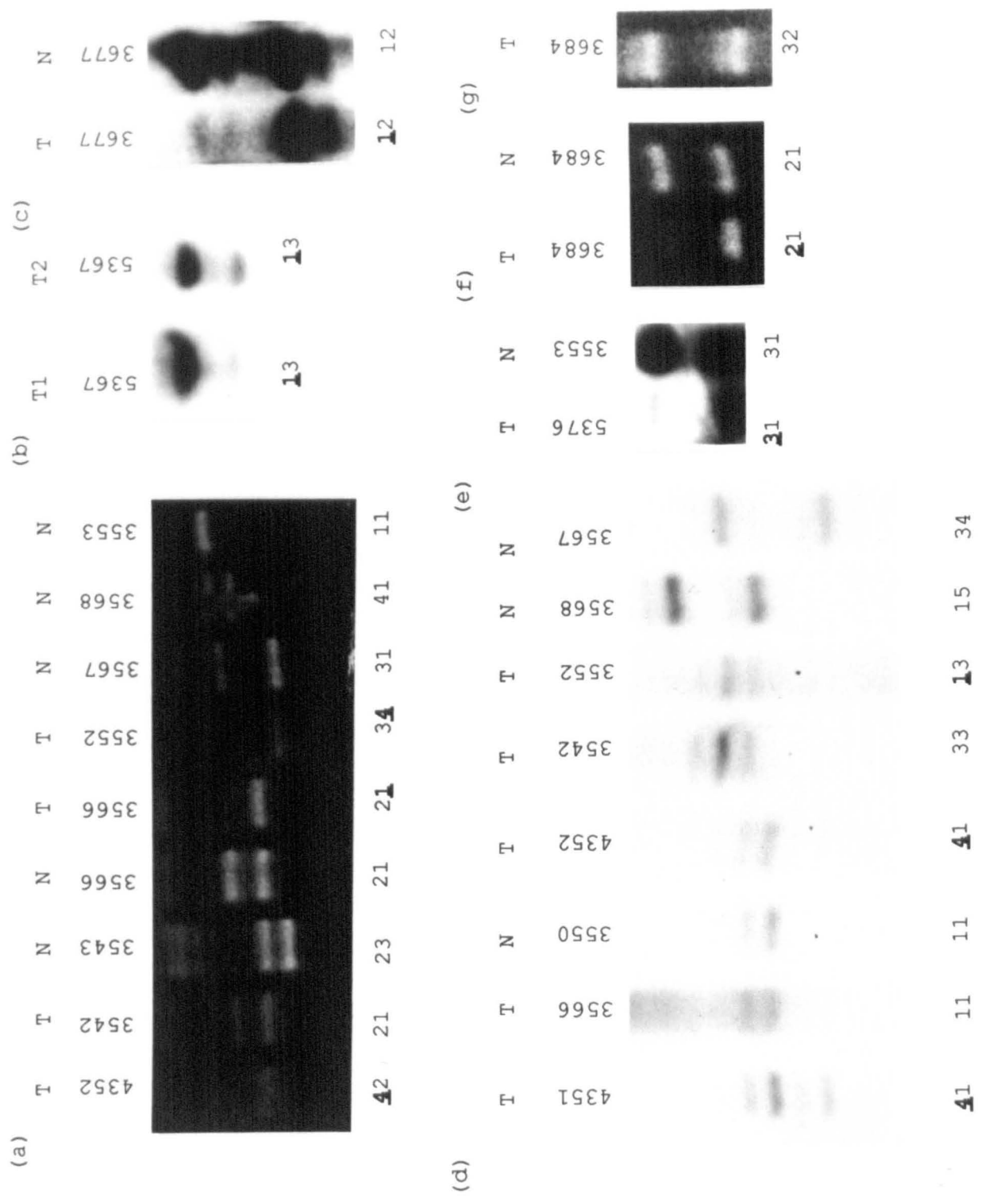
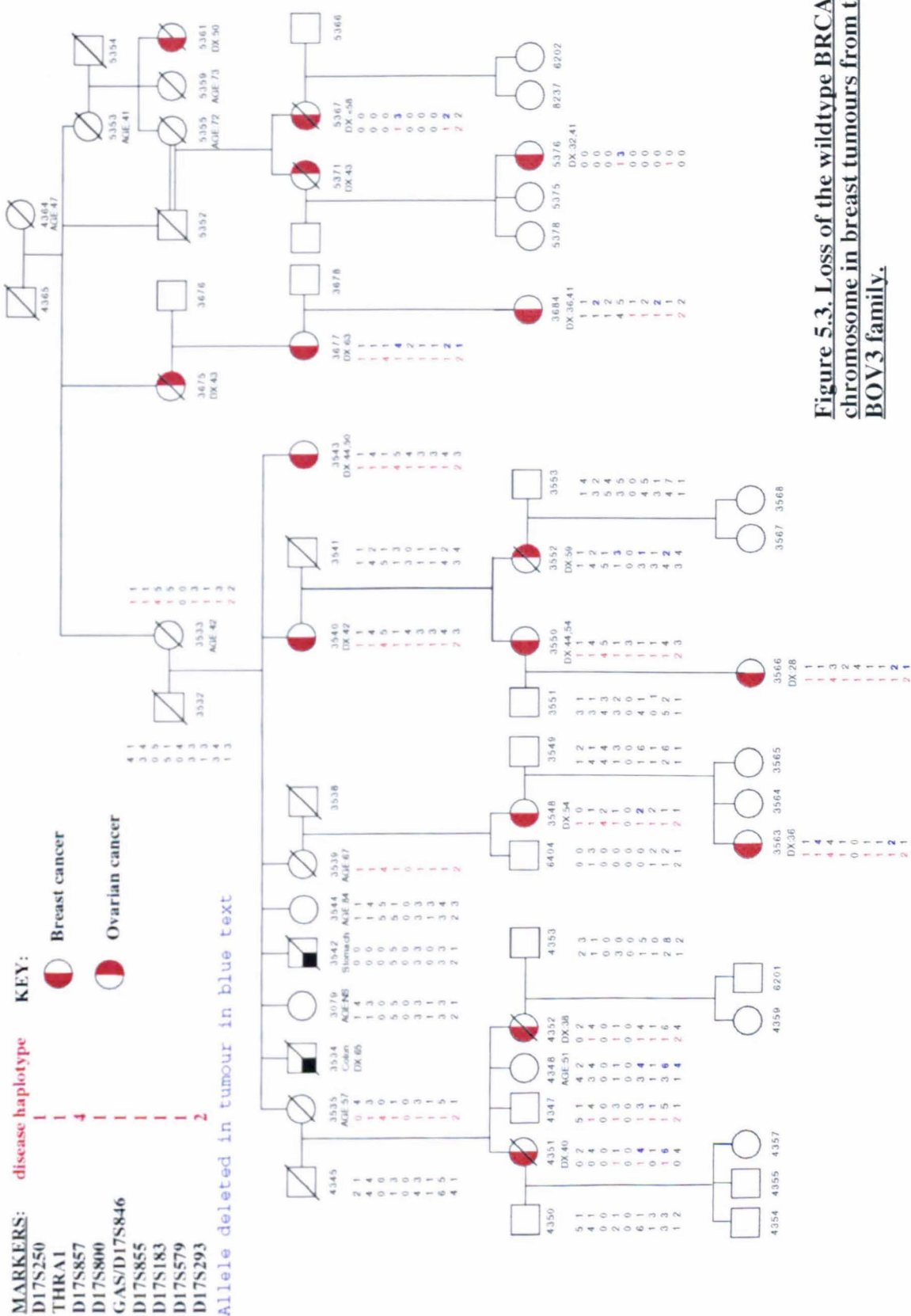


Figure 5.3. Loss of the wildtype BRCA1 chromosome in breast tumours from the BOV3 family.

The pedigree is as described in Figure 3.4. Haplotypes for microsatellite marker loci (as indicated in top left of figure) are shown under each individual typed (as described in Figure 3.4). The haplotype segregating with the disease is indicated in red text (185delAG carrier status was confirmed for all individuals shown to carry the disease haplotype: E. Mavraki, personal communication). Alleles deleted in tumour DNA extracted from archival pathological material are highlighted in blue text: see Figure 5.1 for summary of allele loss data. Loss of the wildtype BRCA1 chromosome was demonstrated in all tumours typed from disease gene carriers. Allele loss was also demonstrated in the DNA extracted from the ovarian tumour of the non-carrier 3552.



In all nine cases, the alleles that are segregating with the disease in the family were retained in the tumours whilst alleles of the chromosome inherited from the non-carrier parent were deleted (Figure 5.3). The resulting hemizyosity for the mutated BRCA1 allele suggests that BRCA1 acts as a tumour suppressor gene, with loss of the wildtype BRCA1 allele a common occurrence in BRCA1-linked breast tumorigenesis (Kelsell *et al*, 1993a). Loss of the wildtype BRCA1 allele has been demonstrated in another study of breast tumours from BRCA1 gene carriers (Smith *et al*, 1992).

The 'knock-out' of the wildtype BRCA1 allele in each tumour involved a deletion of genetic material either side of the BRCA1 region (Figure 5.1). Though the limits of each deletion were not fully characterised, 8 of the carcinomas must have a distal boundary telomeric to BRCA1 with allele loss detected for the microsatellite marker loci: D17S190, D17S293 or D17S588. Deletions for the carcinomas (3563, 3677, 3684, 5367, 5376) extended outside a proximal boundary (D17S800) for BRCA1. Results could not be obtained to define the proximal extent of the deletions on 17q for the other three carcinomas (4352, 4351, 3566) due to uninformativeness of the proximal microsatellite marker loci in these individuals or lack of a PCR amplification product. Tumour DNA from individual 3548 demonstrated allele loss at D17S855 an intragenic BRCA1 marker locus. However, allele loss was not assessed for other microsatellite marker loci.

A plausible target gene other than BRCA1 for this allele loss could be the tumour suppressor gene, TP53 which maps to the short arm of chromosome 17 and in which mutations have been detected in a number of breast carcinomas (see section 1.14.3(iii); for example, Coles *et al*, 1992). To investigate this possibility, a microsatellite marker locus mapping within the TP53 gene (Jones and Nakamura, 1992) was used to determine TP53 allelic status in the carcinomas. Allele loss at the TP53 locus was detected in two of the five informative BRCA1 familial carcinomas analysed (Figure 5.1). This suggests that the focus for allele loss on chromosome 17 in the BOV3 tumours is on the long arm encompassing the BRCA1 locus. This is supported by the

observation that for each BRCA1-associated tumour, loss of the wildtype copy of BRCA1 was demonstrated. An example of retention of both copies of the TP53 locus is shown for the tumour DNA from individual 3684 in Figure 5.2(g). The intensity of the alleles at TP53 in the tumour DNA was comparable to that of the normal lymphocyte DNA from individual 3684 (gel data not shown). Potentially, genes other than BRCA1 on 17q could also be the target for allele loss but, as each deletion spans the wildtype BRCA1 allele, BRCA1 is the likely target.

An additional tumour from the BOV3 pedigree was assessed for allele loss at the BRCA1 locus. This was from the single sporadic ovarian tumour in individual 3552. DNA from this tumour was hemizygous for the BRCA1 region, having retained the maternally derived chromosome. This is consistent with a role for BRCA1 in sporadic ovarian tumorigenesis (Foulkes *et al*, 1993; Saito *et al*, 1993). The possible involvement of BRCA1 in sporadic breast cancer is addressed in section 5.4.

5.3. Allele loss at BRCA1 in familial breast tumours

From analysis of breast tumour DNA from the BOV3 family (previous section) and from other BRCA1-linked tumours (Smith *et al*, 1992), there is strong evidence that BRCA1 acts as a tumour suppressor gene. As loss of the wildtype BRCA1 appears to be a common event in the BRCA1-linked tumours analysed thus far, tumour DNA analyses were conducted in a series of ICRF breast and breast-ovarian cancer families. The purpose of this study was twofold:

- (i) to extend the pedigrees by typing archival tumour material from deceased affected family members for haplotype and linkage analysis
- (ii) to assess allele loss in tumours from putative gene carriers and to use this information to evaluate the likelihood of disease in each family being linked to BRCA1.

Individuals from 63 ICRF breast cancer pedigrees were typed with microsatellite marker loci spanning the BRCA1 locus (alleles detected as described in sections 2.7.1

and 2.7.3, unless otherwise stated) and, depending on tumour sample availability and the generation of PCR products, assessed for allele loss at BRCA1 (as described in section 2.8). All tumour DNA was extracted from archival pathological sections (as described in section 2.3.3). The same set of families were also typed with microsatellite marker loci flanking BRCA2 to determine the proportion of those families likely to be linked to each cancer susceptibility locus (see Chapter 7 for brief analysis of results from this study). Allele loss data was incorporated into the multi-point linkage analysis using the method described in section 2.13.1.2 (Prof. T. Bishop, personal communication). Examples of families analysed with allele loss data are described below.

5.3(i) BC5

This family is classified as a breast-ovarian cancer family (Figure 5.4). The majority of cancers were diagnosed before age 40, with three female family members diagnosed with bilateral breast cancer (295) or breast-ovarian cancer (289, 291). The haplotype (1:1:3) appears to be segregating with the disease. Unaffected individual 287 has inherited only one of the inferred parental chromosomes (constructed on the basis of each haplotype being present in at least two family members). Non-paternity would seem the likeliest interpretation of this result.

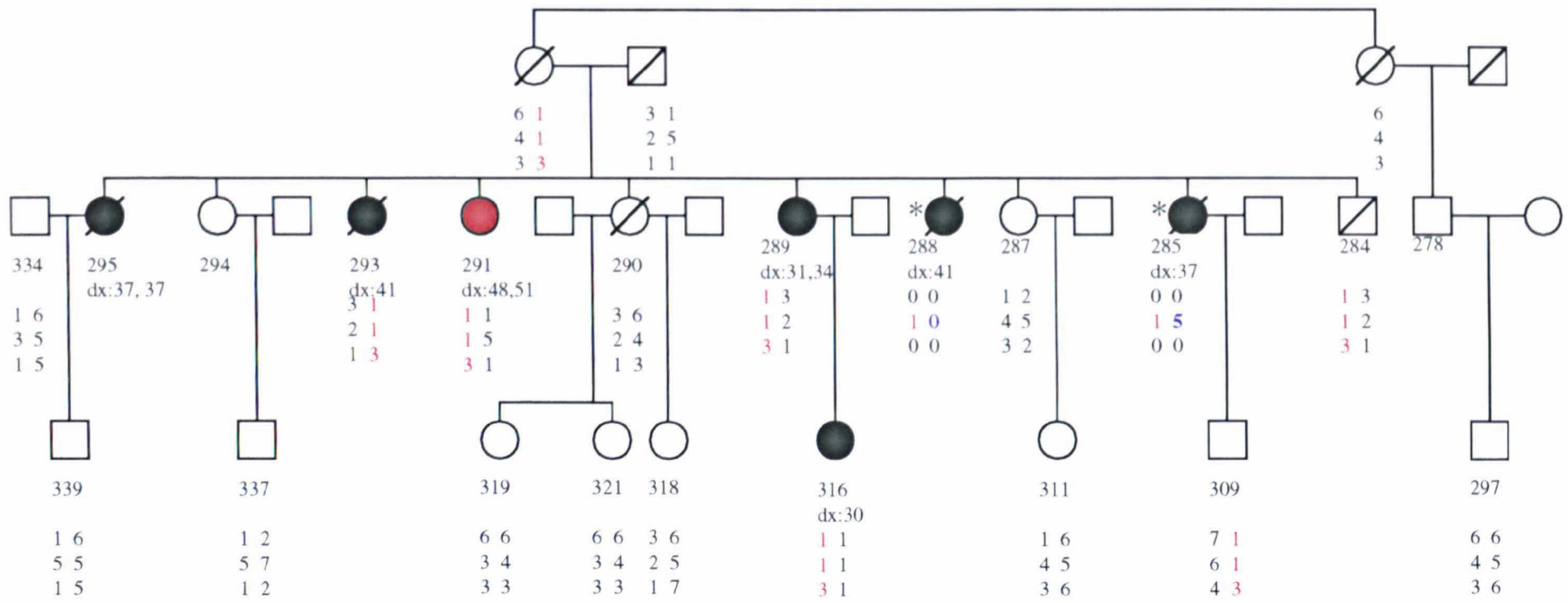
DNA was analysed from two archival pathological breast tumour samples from this family. Tumour DNA from individual 285 revealed loss of the "5" allele at D17S855 with retention of the "1" allele (gel data not shown: interpreted in Figure 5.4). Confirmation for this individual carrying the putative disease haplotype was demonstrated from typing her son's lymphocyte DNA (individual 309). Allele loss at D17S855 was also likely from analysis of tumour DNA from individual 288 (Figure 5.4). Only allele "1" was detected from the PCR amplification of DNA from this tumour at this locus (gel data not shown: interpreted in Figure 5.4). Though a non-tumour DNA sample was unavailable for this individual, from the construction of the inferred parental haplotypes for this individual only one parental haplotype can bear the

Figure 5.4. Haplotype and allele loss analysis of BC5

Symbols are as described in Figure 1.1. The order of microsatellite marker loci and the putative disease haplotype segregating with disease in this family are shown in the bottom left hand corner of the figure. The putative disease haplotype is indicated in red. Haplotypes for individuals 293 and 290 are inferred as are those of their parents. Some haplotypes for individuals at risk of being BRCA1 gene carriers (from pedigree analysis) are not shown as they are of possible clinical relevance. It is likely that the paternal haplotype of individual 287 was inherited from a different father to that of her brother and sisters (a cross-over event between THRA1 and D17S855 is likely to have occurred on the maternal chromosome homologue transmitted to individual 287). This is partly based on the observation of a 6:4:3 haplotype in individual 297 (likely to have been transmitted from the mother of 278 who is the sister of the mother of the large sibship) which is also present in individual 290. In addition, the observation of the haplotypes; 3:2:1 and 1:5:1 present in at least three sibs supports 287 having inherited the haplotype 2:5:2 from a different father. Individuals 288 and 285 were typed from archival breast tumour tissue sections (indicated by an asterisk). The allele deleted in the tumour is indicated in blue text. A zero allele in a tumour DNA sample indicates either unable to generate a PCR amplification product or predicted hemizyosity (at D17S855 in tumour DNA of individual 288). LOD scores supporting disease susceptibility and the inheritance of a mutated BRCA1 gene are shown in the bottom right of figure. A frameshift mutation in one allele of BRCA1 (deletion of nucleotide T at position 3896: 3896delT) has been detected in lymphocyte DNA extracted from individuals: 291, 289 and 316: Xu *et al*, in press).

Figure 5.4. Haplotype and allele loss analysis of BC5

132



Marker	Disease Haplotype
THRA1	1
D17S855	1
D17S579	3

LOD scores at BRCA1:
without allele loss data: 1.25
with allele loss data: 1.41
* typed from tumour DNA

"1" allele at D17S855. This suggests tumour DNA from individual 288 is hemizygous at D17S855 with the same allele retained as that detected in the tumour DNA from 285. Haplotype data from her offspring confirmed this analysis (data not shown as clinically relevant).

Analysis of these two tumours and the segregation of a 17q12-21 haplotype with the disease suggests cancer predisposition in this family is due to an inherited mutation in BRCA1. A LOD score at BRCA1 of 1.41 incorporating the allele loss data (Prof. T. Bishop, personal communication) supports disease in this family being linked to BRCA1 (Figure 5.4). Subsequently, a frameshift mutation in one allele of BRCA1 (deletion of nucleotide T at position 3896; 3896delT) has been detected in the germline DNA from gene carriers (identified from haplotype analysis: this section) in this family (Xu *et al*, in press).

5.3.(ii) BOV4

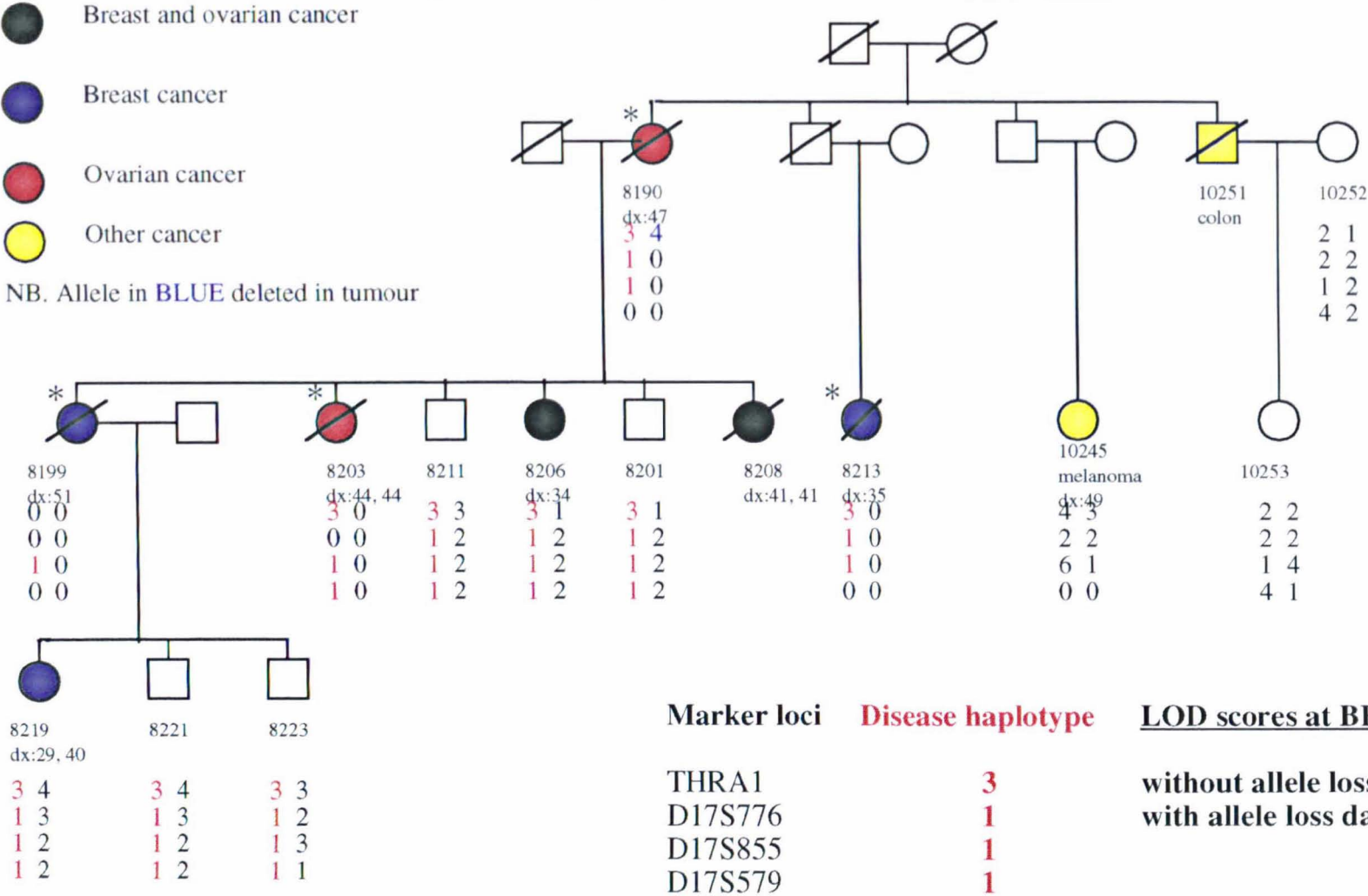
This pedigree is classified as a breast-ovarian cancer family with seven female members affected with either breast or ovarian cancer, or both (Figure 5.5). The majority of affected members in this family are deceased. However archival tumour DNA was available from four affected women in this family: 8199, 8203, 8190 and 8213. Microsatellite analysis of DNA extracted from these tumours showed only a single allele at each locus tested (for example, D17S855). This could be explained by either homozygosity in the germline or hemizygous loss in the tumour DNA at BRCA1. A signal of reduced intensity for allele "4" was detected in the ovarian tumour DNA from individual 8190 at the THRA1 locus indicating a reduction to hemizygosity with retention of allele 3 (Figure 5.6).

Though the constitutive allelic status of the tumours could not be confirmed (due to lack of a "normal" DNA sample), the observation that all the affected members share a common haplotype (including that detected in the tumour DNA) would support disease in this family being linked to BRCA1. A LOD score at BRCA1 of 1.43 incorporating

Figure 5.5. Haplotype and allele loss analysis of BOV4

Symbols are as described in Figure 1.1. Key indicating the cancer type for affected individuals is indicated in the top left hand corner of the figure. The specific cancer type for those individuals affected with cancer other than breast and/or ovarian cancer is indicated under the individual's identification number with the age of diagnosis (dx). The order of microsatellite marker loci and the putative disease haplotype segregating with disease in this family are shown in the bottom right hand corner of the figure. The putative disease haplotype is indicated in red. Individuals 8199, 8203, 8190 and 8213 were typed from archival breast tumour tissue sections (indicated by an asterisk). Allele deleted in tumour is indicated in blue text. A zero allele in a tumour DNA sample indicates either unable to generate a PCR amplification product, or unable to determine if tumour DNA is hemizygous or constitutively homozygous at that locus. LOD scores (particularly that generated from the incorporation of allele loss data) supporting disease susceptibility and the inheritance of a mutated BRCA1 gene are shown at the bottom right of the figure. A frameshift mutation in one allele of BRCA1 (deletion of four nucleotides at position 3875: 3875del4) has been detected in lymphocyte DNA extracted from individuals 8219 and 8206 (Xu *et al*, in press). In addition, individual 8219 has inherited a BRCA1 mutation in the paternally transmitted BRCA1 allele (Dr. D. Black, personal communication).

Figure 5.5. Haplotyping and allele loss analysis of BOV4



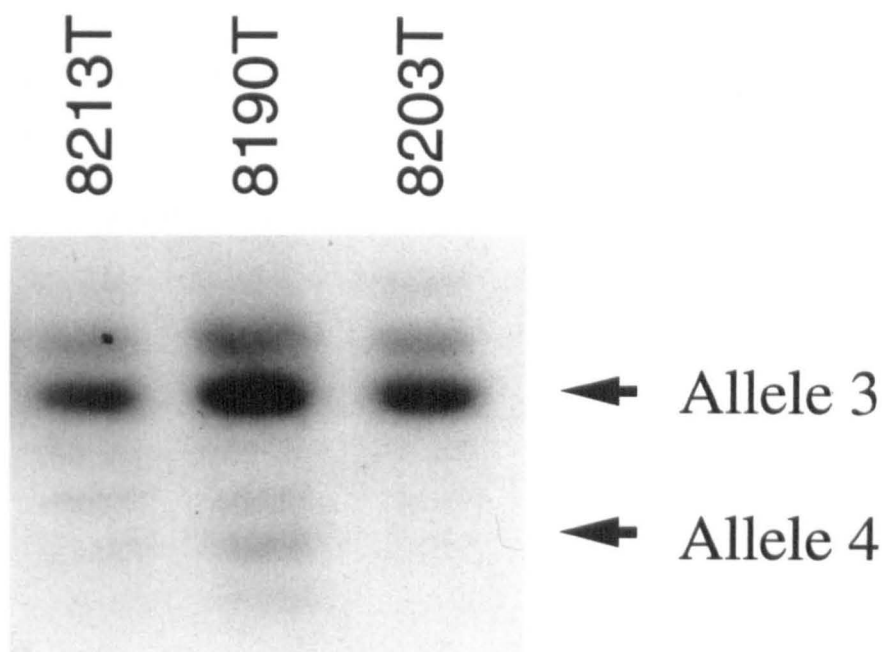


Figure 5.6. Allele loss at the THRA1 locus in ovarian tumour DNA of individual 8190 from BOV4

Examples of the results of microsatellite marker loci analysis of tumour DNA from the BOV4 pedigree. The personal identification number is indicated above each lane (further individual details such as sex, disease and gene carrier status can be interpreted from Figure 5.5). DNA was extracted from pathological material with approximately 70-90 % tumour cell content (estimated as described in section 2.2.5). Allele numbering is shown to the right of the figure. Allele 3 at the THRA1 locus is segregating with breast and ovarian cancer (see Figure 5.5). All three samples have inherited allele 3. Loss of allele 4 (from comparison of allele 3 and 4 intensities) is clearly evident in the ovarian tumour DNA extracted from individual 8190. The loss of the wildtype allele in this tumour supports BRCA1 acting as a familial tumour suppressor gene in BRCA1-associated ovarian tumorigenesis. As non-tumour DNA was unavailable for analysis for the individuals 8213 and 8203, allele loss in their respective tumour DNA could not be determined as both individuals may be constitutively homozygous at the THRA1 locus. However, it is possible that allele 4 is also deleted in the tumour DNA from individual 8213.

the allele loss data (Prof. T. Bishop, personal communication) further suggests that disease in this family is linked to BRCA1 (Figure 5.5). Subsequently, a frameshift mutation (3875del4) in one copy of the BRCA1 gene has been detected in the two family members from this pedigree for which lymphocyte DNA was available (8206, 8219: Xu *et al*, in press). Interestingly, a BRCA1 mutation (4148G: Arginine to Glycine amino acid change) has also been detected in the paternally inherited BRCA1 allele in individual 8219 (Dr. D. Black, personal communication). This compound heterozygote is the second reported case of an individual with two putative mutant BRCA1 alleles. The other was detected in a woman diagnosed with breast cancer at age 32 who was homozygous for an AA deletion at position 2800 (Boyd *et al*, 1995). Both women appear normal and developed breast cancer at a comparable age to BRCA1 heterozygotes from each pedigree. The biological significance of carrying two abnormal copies of BRCA1 requires investigation and is discussed further in Chapter 7.

These results (including those generated from the analysis of the BOV3 tumour DNA) were combined with data generated by other members of the Breast Cancer Consortium Linkage Group. 86% of BRCA1-linked breast-ovarian carcinomas showed loss of the BRCA1 region with the disease haplotype retained in each case (Cornelis *et al*, 1995). This analysis provided firm evidence that BRCA1 acts as a tumour suppressor gene and deletion of the wildtype allele is a common mechanism of inactivation in breast and ovarian carcinomas from BRCA1 gene carriers. Additionally, there was no indication of an imprinting effect as chromosomal losses involved both paternally and maternally derived chromosomes (Cornelis *et al*, 1995). An imprinting effect could not be excluded as a possibility from the allele loss study of breast tumours from the BOV3 pedigree (Figure 5.3) as all tumours from gene carriers analysed had inherited the mutant BRCA1 allele from their mothers.

A selection of the other ICRF pedigrees analysed by incorporating allele loss data are described in the following sections:

Figure 5.7. BRCA1 haplotype and allele loss analysis in three ICRF pedigrees

The pedigrees are constructed as described for other pedigrees, for example, Figure 3.4. Affected status of family members is as described in the key (in bottom left of figure). The cancer type is indicated under the identification number for each affected member: BC=breast cancer; OV=ovarian cancer; BOV=breast and ovarian cancer. An asterisk * to the left of an individual's symbol indicates genotyping and allele loss assessment of tumour DNA was performed. For individuals 6435 (Figure 5.7(i)), 24178 and 24179 (Figure 5.7(ii)), lymphocyte DNA was also available for genotyping and comparison of allele status to that obtained from tumour analysis. Haplotypes for microsatellite marker loci (as indicated to the left of typed individual in last generation of each family) are shown under each individual typed. The haplotype putatively segregating (in pedigrees BOV178 and BC2148) with the disease is indicated in red text. Alleles deleted in tumour DNA extracted from archival pathological material are highlighted in blue text. Loss was detected in each tumour for each locus analysed or for which a PCR amplification product was generated. Linkage analysis was conducted as described in section 2.13.1.2.

(i) BOV178

Allele loss was detected in both tumours analysed with the retention of the same haplotype: 1: 2: 3: 3. This haplotype appears to be segregating with the disease in the family. Incorporation of allele loss data into the linkage analysis supports the implication that disease in this family is due to a mutation in BRCA1 (see LOD scores under the BOV178 pedigree). Subsequently, a 185delAG mutation has been detected in one allele from individuals 6432 (diagnosed with breast and ovarian cancer) and 16919 (unaffected male gene carrier) (Xu *et al*, in press).

(ii) BC2148

Allele loss was detected in both tumours analysed with the retention of the same haplotype: 1: 1: 1. This haplotype appears to be segregating with the disease in the

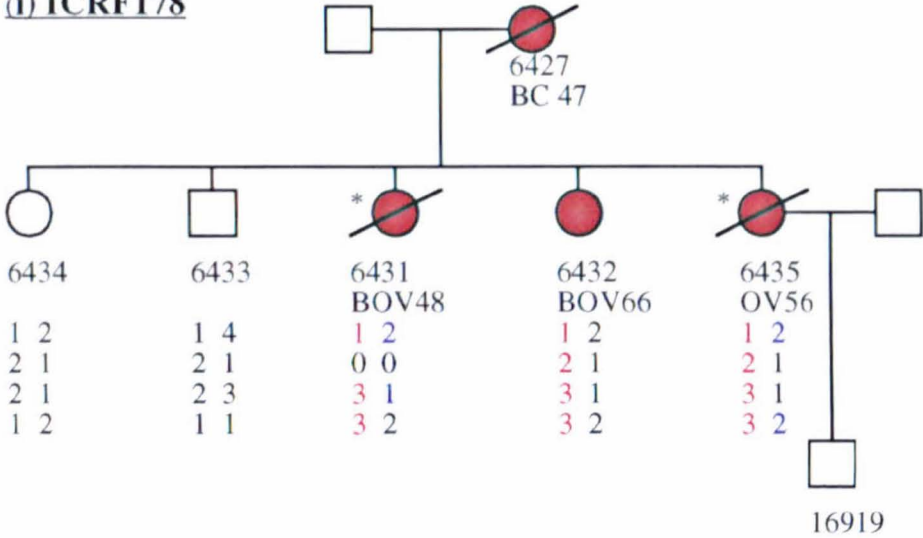
family. Incorporation of allele loss data into the linkage analysis supports the implication that disease in this family is due to a mutation in BRCA1 (see LOD scores to the left of the BC2148 pedigree). No BRCA1 mutation has been reported in this family as yet (Prof. E. Solomon, personal communication).

(iii) BC1777

Allele loss was detected at all three microsatellite marker loci analysed from the breast tumour of individual 24449. However, the haplotype retained (1: 4: 1) was not the haplotype in common with the two other genotyped affected pedigree members. This suggests disease susceptibility in this family is not due to an inherited mutation in BRCA1. Linkage analysis has yet to be conducted in this family (Prof. T. Bishop, personal communication). No BRCA1 mutation has been reported in this family (Prof. E. Solomon, personal communication).

Figure 5.7 BRCA1 haplotype and allele loss analysis in three ICRF pedigrees

(i) ICRF178



LOD score without allele loss: 0.59
LOD score with allele loss: 0.99

THRA1 1 3
D17S800 2 2
D17S776 3 3
D17S579 3 4

KEY:

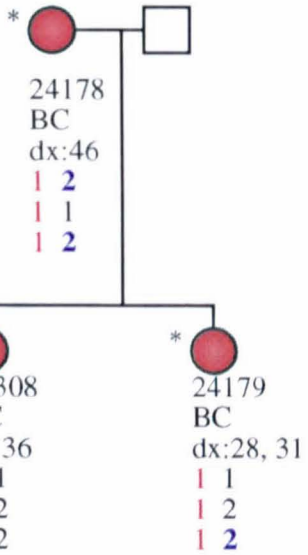
breast or ovarian cancer

Putative disease haplotype in **red**

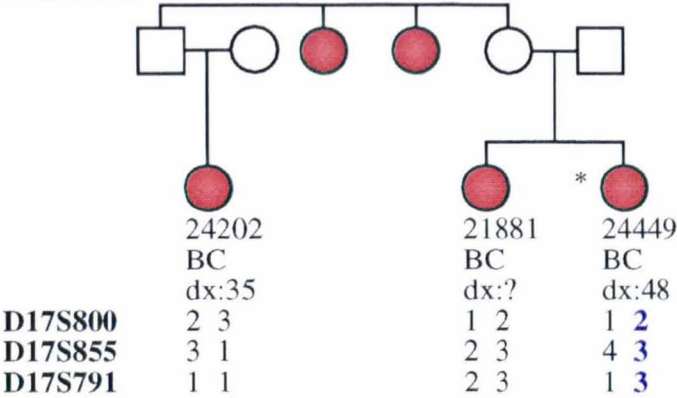
* tumour analysed for LOH, alleles deleted in **blue**

(ii) BC2148

LOD score without allele loss: 0.26
LOD score with allele loss: 0.51



(iii) BC1777



5.3.(iii) ICRF 178

This pedigree consists of three affected sisters; two who developed both breast-ovarian cancer (6431 and 6432) and the third, who developed ovarian cancer (6435: Figure 5.7(a)). Lymphocyte DNA was available from two of the affected sisters (6432, 6435) and a number of unaffected family members. Breast and ovarian tumour DNA from archival pathological material was available for the affected family members: 6431 and 6435, respectively. From comparison of the three affected sister's haplotypes, all three have inherited the same maternal and paternal chromosomes (if a recombination event occurred between BRCA1 and D17S579 in 6431). From pedigree analysis, it is likely that cancer susceptibility was inherited from their mother, who had breast cancer diagnosed at age 47. DNA was unavailable from their parents to determine which haplotype was inherited from each parent. However, DNA from the two tumours analysed revealed allele loss at BRCA1 with retention of the same alleles suggesting the haplotype (1;2;3;3) was segregating with disease in this family. These results were combined into the linkage analysis, generating a LOD score of 0.99 at 0.00 recombination fraction with BRCA1 and the disease (Prof. T. Bishop, personal communication). Subsequent to this analysis, a 185delAG mutation in BRCA1 was detected in this family (Xu *et al*, in press).

5.3.(iv). Ancestral mutation?

The ICRF178 BRCA1 mutation is identical to that seen in the BOV3 family (Chapter 3) and in a third family, ICRF546 (Xu *et al*, in press). Disease haplotype comparison of gene carriers from these three ICRF pedigrees was performed using microsatellite marker loci intragenic to BRCA1 (detected as described in section 2.7.3). As DNA was only available from one gene carrier from ICRF546 and ICRF178 (at time of this analysis), haplotype construction was impossible to infer if the microsatellite marker loci tested were heterozygous in each gene carrier. However, a common disease haplotype could be inferred (by the observation of alleles in common at each locus in gene carriers from each family) suggesting that disease in all three families is due to the same ancestral mutation (Table 5.1). Interestingly, an ancestral 185delAG BRCA1

mutation has been demonstrated to segregate in the Ashkenazi population at a frequency of approximately 1% (Streuwing *et al*, 1995a). However from gene carrier haplotype analysis, the ICRF 185delAG mutation arose independently to that found in the Ashkenazi population (Neuhausen *et al*, 1996a).

The haplotypes from the ICRF gene carriers were extended by genotyping with microsatellite marker loci centromeric and proximal to BRCA1 (Table 5.1). Shared alleles were present in all gene carriers over a region spanning at least 23-25 cM. As the putative shared disease haplotype is extensive, this would suggest that the BRCA1 mutation is of a similar age in all three families, having only very recently diverged. However, interviewed individuals from each family are unaware of any genealogical link between the three families (Prof. T. Bishop, personal communication). There is a divergence of the BOV3 disease haplotype from the other two, centromeric to BRCA1 at D17S925 and telomeric to BRCA1 at D17S787. The two gene carriers from ICRF546 and ICRF178 have a possible disease haplotype in common which spans at least 30 cM (diverge at D17S948) which suggests they may be more closely related to each other than to the BOV3 family. Indeed, a common family name has been identified in earlier generations of the ICRF546 and ICRF178 families supporting this observation (Prof. T. Bishop, personal communication). However, no genealogical link has been confirmed as yet.

5.3.(iv) BC 2148

In this family, the two affected daughters (24308 and 24197) share the same haplotypes (Figure 5.7(b)). Breast tumour DNA was available from the mother, 24178, and the daughter, 24179, and both were assessed for allele loss at BRCA1. Loss was detected in both tumours with the same sized alleles retained in each tumour. This tentatively supports linkage to BRCA1 and is reflected in the increase of the LOD score following incorporation of the allele loss data (Prof. T. Bishop, personal communication).

Table 5.1. Haplotype comparison of the three ICRF breast-ovarian cancer families carrying the 185delAG mutation in BRCA1

Marker locus	ICRF546	ICRF178	BOV3
D17S953 (7 cM)	123	123/129	129
D17S925 (6 cM)	165/172	165	165
D17S798 (5 cM)	218/225	225/227	225
D17S933 (3 cM)	189/191	189/197	189
THRA1 (1-2 cM)	167	167/177	167
D17S776 (0-1 cM)	122	122	122
BRCA1(185delAG)	126	126	126
D17S1327	135	135/153	135
D17S1323	145/149	145/149	145
D17S1322	128	128	128
D17S855	153	151/153	153
D17S1326 (1-2 cM)	101	101/109	101
D17S579 (3 cM)	120/128	120/114	120
D17S806	187	187	187
D17S1827 (3 cM)	128/145	145	145
D17S1820 (4 cM)	265/277	267/277	277
D17S787 (13 cM)	139/153	137/153	141
D17S948	143/151	133/145	143

One gene carrier was typed from ICRF546 (13523: pedigree not shown) and ICRF178 (16919; see pedigree in Figure 5.7(i)) and two gene carriers from BOV3 (3540 and 3550: see pedigree in Figure 3.4). Allele detection (and sizing) was as described in section 2.7.3. Size of PCR amplification DNA products shown in base pairs (bp). For microsatellite marker loci in which the disease allele could not be determined, both alleles are shown (if heterozygous at that locus). Intragenic (or closely flanking) BRCA1 marker loci are indicated in bold text. Approximate genetic distances in centiMorgans (cM) between marker loci is shown in brackets (from Gyapay *et al*, 1994; Albertsen *et al*, 1994; Dib *et al*, 1996). There is a divergence of the BOV3 disease haplotype from the other two, centromeric to BRCA1 at D17S925 and telomeric to BRCA1 at D17S787.

5.3.(v) BC 1777

Lymphocyte DNA was available from three female members of this family who have developed breast cancer (24202, 21881 and 24449: Figure 5.7(c)). Analysis suggested a putative common haplotype in the BRCA1 region from D17S800 to D17S855, though other haplotypes could be constructed. Allele loss was detected in breast tumour DNA from individual 24449. However, the alleles retained (1:4:1) were not from the putative "shared" haplotype (2:3:3/1). This would indicate that breast cancer in this family is not linked to BRCA1, or, that individual 24449 has "sporadic" disease (or has inherited a distinct cancer susceptibility mutation). Linkage analysis has yet to be performed in this pedigree (Prof. T. Bishop, personal communication).

5.3.(vi) ICRF 510

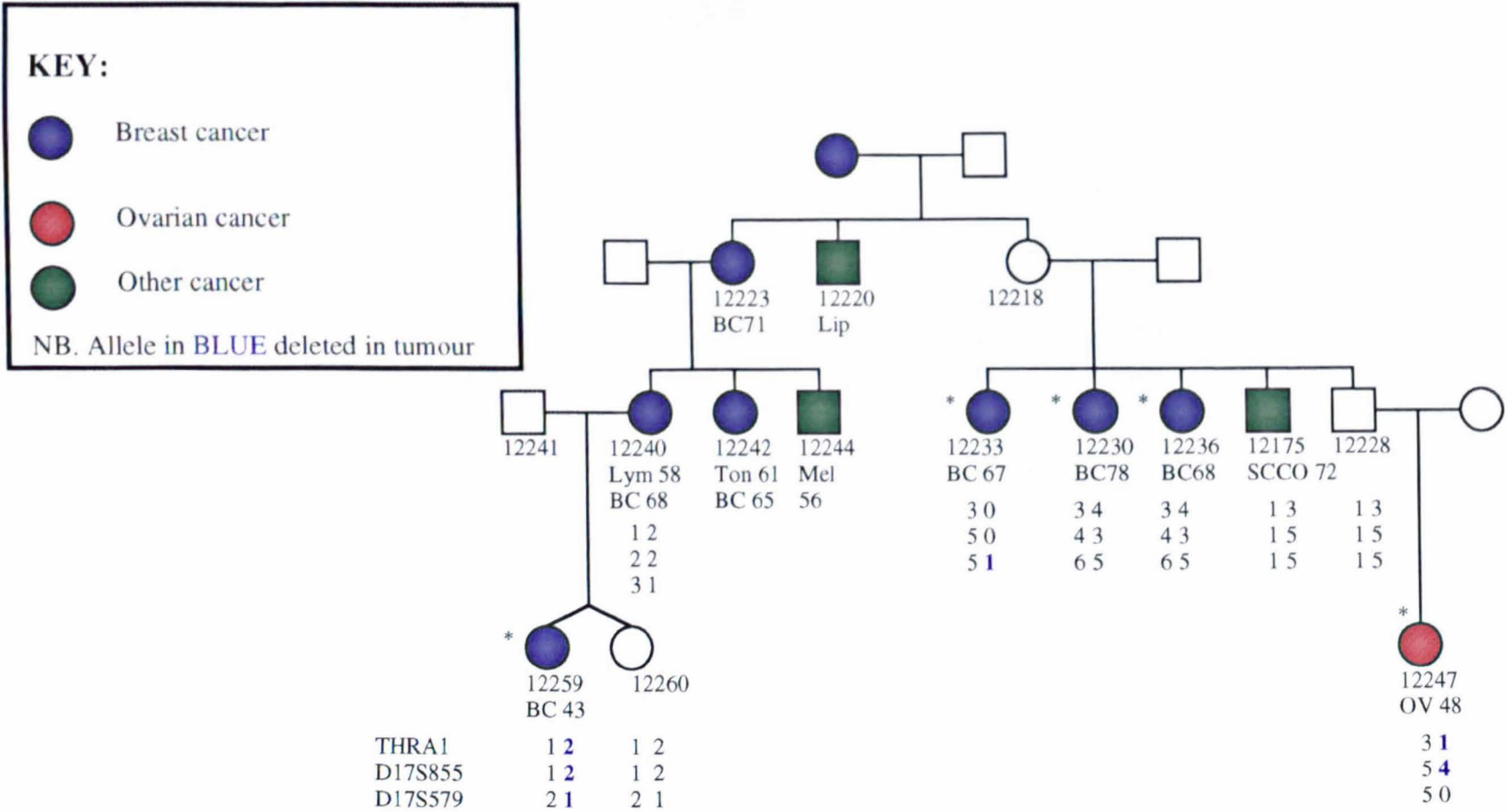
This pedigree is a multi-site cancer family, including five confirmed diagnosed cases of breast cancer and one case of ovarian cancer. There is no common BRCA1 haplotype shared in the majority of the breast cancer cases or with the ovarian cancer case (Figure 5.8). This would indicate that disease susceptibility in this family is not linked to BRCA1, reflected in a LOD score of -0.44 for BRCA1 and the disease (Prof. T. Bishop, personal communication). Analysis of tumour DNA from family members supports this finding. No allele loss at BRCA1 was detected in breast tumours from individuals 12230 and 12236. Other compelling evidence comes from the analysis of a breast tumour from individual 12259. Allele loss was detected at the BRCA1 locus (Figure 5.9), but the allele retained was from the paternal chromosome (Figure 5.8). From pedigree analysis, the putative cancer susceptibility locus appears to be inherited from her mother.

Microsatellite marker loci mapping to other candidate regions which harbour cancer susceptibility genes, for example, TP53, have been screened across lymphocyte and tumour DNA samples from this family. No indication of linkage with the disease to these loci or common patterns of allele loss have been found (data not shown).

Figure 5.8. ICRF510 BRCA1 haplotype and tumour DNA analysis

Symbols are as described for previous pedigrees, for example, Figure 3.2. The key indicating the cancer type for affected individuals is indicated in the top left hand corner of the figure. The specific cancer type for those individuals affected with cancer is under the individual's identification number: BC=breast cancer, OV=ovarian cancer, SCCO=squamous carcinoma of the oesophagus, Lym=lymphoma, Ton=tongue carcinoma, Lip=lip carcinoma, Mel=melanoma. The order of microsatellite marker loci is shown in the bottom left hand corner of the figure. Haplotypes are shown under the identification number for each individual genotyped. For individuals 12259, 12233, 12230, 12236 and 12247, archival breast tumour DNA was analysed (indicated by an asterisk*). However, for individuals: 12233 and 12247, tumour DNA was only available for genotyping. A zero allele indicates they are either hemizygous or homozygous at that locus (as genotyped from tumour DNA with no normal DNA sample available). Allele deleted in tumour is indicated in blue text. Allele loss was detected in three of the tumours analysed. For tumours: 12233 and 12247, the same haplotype (3: 5: 5) was retained in each tumour. However, for breast tumour DNA derived from 12259, the haplotype retained was inherited from the non-disease transmitting parent (from pedigree analysis). Example of allele loss in the tumour DNA from this individual is shown in Figure 5.9. Haplotype and allele loss data suggests disease in this family is not due to an inherited mutation in BRCA1 (supported by a LOD score of -0.44 at BRCA1 with the disease in this family: Prof. T. Bishop, personal communication).

Figure 5.8. ICRF510 BRCA1 haplotype and tumour DNA analysis



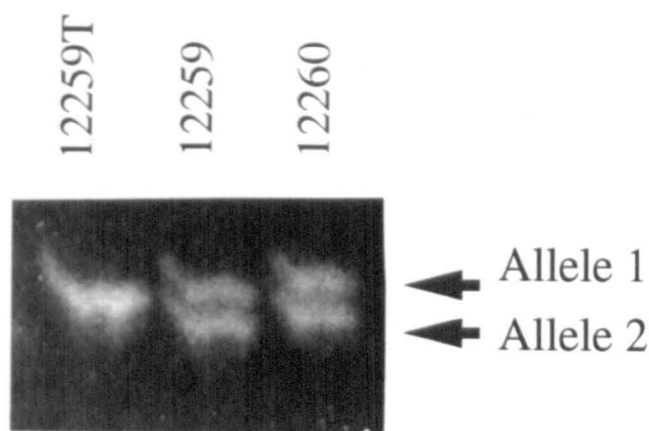


Figure 5.9. Allele loss at the THRA1 locus in breast tumour DNA of individual 12259 from ICRF510

The personal identification number is indicated above each lane (further individual details such as sex, disease and gene carrier status can be interpreted from Figure 5.8). Allele status of breast tumour DNA from individual 12259 at THRA1 is shown in the lane indicated by the label, 12259T. The other two lanes show allele status at the same locus in lymphocyte DNA from the same individual 12259 and her twin sister, 12260. Tumour DNA was extracted from pathological material with approximately 80 % tumour cell content (estimated as described in section 2.2.5). Allele numbering is shown to the right of the figure. Allele detection is as described in section 2.7.1 and allele loss interpreted as described in section 2.8. Loss of allele 1 (from comparison of allele status in DNA from lymphocyte DNA of the same individual) is clearly evident. The allele retained (allele 1) in the tumour DNA was likely to have been inherited from her father (based on haplotype construction for that branch of the family: see Figure 5.8). From pedigree analysis, disease susceptibility appears to be transmitted from her affected mother, 12240.

5.4. BRCA1 allele loss in sporadic breast carcinomas

118 paired tumour and lymphocyte DNA samples were analysed from a sporadic female breast cancer population (individuals who have no apparent family history) in the south east of England. No selection for histology or age of onset of cancer was used in sampling. The DNA analysed was extracted from fresh tumour material and lymphocytes (as described in section 2.2.4). These breast tumour/blood DNA pairs were screened with microsatellite marker loci flanking BRCA1 (THRA1, D17S800, D17S855, D17S579: see Figure 3.1). Analysis of the BRCA1 status in these carcinomas might indicate a possible role for this gene in sporadic disease. All microsatellite allele profiles were detected and analysed as described in sections 2.7.3 and 2.8, respectively. Allele loss at BRCA1 was identified as complete or almost complete reduction (for the majority of cases, a 20-50% reduction was observed) to hemizyosity at the closest informative loci to BRCA1(at least two loci).

47 % of carcinomas (56/118) were hemizygous at BRCA1. Background or random allele loss rates are estimated to be seen in between 10 and 15 % of tumours, implying that loss of the BRCA1 chromosomal region is a common event in sporadic breast tumorigenesis. This has been demonstrated by a number of other studies (for example, Futreal *et al*, 1992). However since the identification of the BRCA1 gene, a second somatic mutation event in the retained allele from sporadic breast carcinomas demonstrating allele loss at this locus has not been identified (Futreal *et al*, 1994; reviewed Szabo and King, 1995). Therefore, the role of BRCA1 as a classical tumour suppressor gene in sporadic disease is unclear.

It has been proposed that there are other target regions for allele loss in breast (and ovarian) tumours on the long arm of chromosome 17 which are distinct from BRCA1 (Cornelis *et al*, 1993; Saito *et al*, 1993; Jacobs *et al*, 1993; Radford *et al*, 1993). One of these proposed target regions for allele loss is 17q24-25. Twenty of the 118 sporadic carcinomas were examined for allele loss in this region (detected as described in section 2.7.3). Of interest to this study is that the TOC locus associated with

oesophageal cancer susceptibility also maps in this region (Chapter 6). These breast tumour/lymphocyte DNA pairs were screened with microsatellite marker loci mapping to the TOC region (D17S929 and D17S1603: see Figure 6.5). The results are summarised in Table 5.2.

Table 5.2. 17q allele loss in sporadic breast tumours

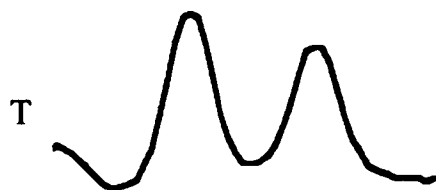
Chromosomal Region	Number of tumours
Loss at BRCA1 and 17q24-25	10
Retained at BRCA1 and 17q24-25	6
Loss at BRCA1 only	3
Loss at 17q24-25 only	1

A number of the tumour DNA samples analysed demonstrated specific sites of deletion on 17q. Three breast carcinomas demonstrated allele loss at BRCA1 but not at 17q24-25, suggesting that the target tumour suppressor gene may be BRCA1. One tumour (1133) demonstrated allele loss at 17q24-25 but not at BRCA1 (Figure 5.10). However, the number of tumours analysed in this study are too small to support the presence of a tumour suppressor gene telomeric to BRCA1. Indeed, from analysis of microsatellite marker loci mapping to the region harbouring TP53 at 17p13 in the panel of 118 breast tumour/blood pairs, allele loss was detected in the vast majority of tumours also displaying allele loss on 17q (data not shown and V. Murray, personal communication). This suggests loss of the whole of one copy of chromosome 17 may be a common mechanism in breast tumorigenesis.

5.5. BRCA2 allele loss in sporadic breast carcinomas

Studies have shown that like BRCA1, BRCA2 also acts as a tumour suppressor gene with characteristic loss of the wildtype chromosome in tumours from BRCA2 gene carriers (Collins *et al* 1995; Thorlacious *et al*, 1995; Gudmundsson *et al*, 1995). The 118 paired breast tumour and lymphocyte DNA samples, previously assessed for BRCA1 allele loss (section 5.3), were screened with microsatellite marker loci

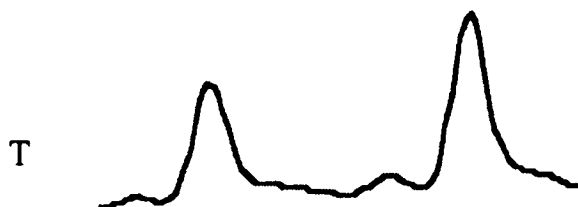
(a) D17S579



154

156

(b) D17S929



222

228

Figure 5.10. Pattern of allele loss on 17q in DNA from a “sporadic” breast tumour 1133

Allele status in tumour DNA from 1133 at two microsatellite loci on the long arm of chromosome 17. In this tumour, loss of the region 17q24-25 (D17S929) is demonstrated, but there was no evidence of allele loss at 17q12-21 (D17S579). D17S579 maps approximately 1-2 cM distal to BRCA1 (Albertsen *et al*, 1995). The alleles were detected (and sized) as described in section 2.7.3 and allele status interpreted as described in section 2.8. At each locus, the tumour DNA profile (T) is shown above that obtained from the lymphocyte DNA (N) of the same individual. Allele sizes (bp) are shown underneath each allele peak.

- (a) D17S579. Identical allele profile in the tumour DNA compared to that in the lymphocyte DNA indicates no allele loss in the tumour at this locus.
(b) D17S929. Reduction in the relative peak height of allele 222 in the tumour DNA compared to that in the lymphocyte DNA indicates loss of this allele in the tumour.

flanking BRCA2 (D13S260, D13S171, D13S267 (Figure 4.2)). Allele profiles were detected and assessed as described in sections 2.7.3 and 2.8.

At BRCA2, 45 % of breast carcinomas (53/118) were found to demonstrate allele loss. This would suggest that deletion of BRCA2 is also a common mechanism in breast carcinogenesis. Genes other than BRCA2 mapping to 13q could be the target for this deletion event, for example, the retinoblastoma gene (RB1). However a number of breast tumours have been demonstrated to show specific loss of the BRCA2 region and not at RB1 (for example: Gudmundsson *et al*, 1995; Cleton-Jansen *et al*, 1995).

Recent studies have demonstrated that somatic DNA mutations in the BRCA2 gene in breast tumours are, like BRCA1, extremely rare (Lancaster *et al*, 1996; Teng *et al*, 1996; Miki *et al*, 1996).

5.6. Allele loss of BRCA1/ BRCA2 in grade 3 breast carcinomas

Breast tumour pathological material from seven affected members of the BOV3 pedigree (see section 5.2) known to carry a BRCA1 mutation was analysed and all were found to be grade 3 infiltrating ductal carcinomas (Gusterson *et al*, 1996). In contrast, the histology of the ovarian tumour from a non-gene carrier in the family (individual 3552: section 3.5(a)) was described as a well to moderately differentiated tumour (Dr. B. Gusterson, personal communication). Other observations support a significant association of grade 3 infiltrating ductal carcinoma of the breast with female BRCA1 gene carriers (Jacquemier *et al*, 1995; Eisinger *et al*, 1996). Preliminary histological analysis of breast carcinomas from affected individuals from ICRF families likely to be BRCA2 gene carriers (from haplotype, linkage and, for one family, mutation analysis) found no association with any particular histological grade (Prof. T. Bishop, personal communication).

By studying somatic events in phenotypically similar tumours with similar genetic susceptibility backgrounds, it may be possible to identify common pathways of

somatic alterations and identify the genes mutated during tumour progression. This section describes an investigation into the relationship between BRCA1 and BRCA2 in the carcinogenic process and, in particular, to determine whether BRCA1 and BRCA2 are involved in the same or separate tumour development pathways. The initial focus were seven grade 3 carcinomas from the BOV3 pedigree which all carry the same BRCA1 mutation (Chapter 3). Previous tumour DNA analysis in the seven breast carcinomas had shown that all demonstrated loss of the chromosome harbouring the wildtype copy of BRCA1 (section 5.2). These tumour DNA samples were screened with microsatellite marker loci flanking BRCA2 (D13S260, D13S171, SLS234, D13S267: see Figure 4.1) to assess the extent of allele loss at this tumour suppressor gene locus (detected as described in 2.7.3).

Interestingly, allele loss was clearly observed in five BOV3 carcinomas with at least two informative markers flanking the minimal interval for BRCA2 (Table 5.3 and Figure 5.11 (a)). The other two carcinomas: 4351 and 4352 gave equivocal evidence (due to lack of a normal DNA control for these individuals; allele profiles were compared to those of their offspring). This result suggests that somatic inactivation of at least one copy of BRCA2 may have an important role in the development of grade 3 tumours derived from BRCA1 gene carriers. DNA mutation analysis in the retained BRCA2 allele has not yet been performed in the BOV3 tumours.

Two other chromosomal regions putatively harbouring tumour suppressor genes were also examined and the results are also shown in Table 5.3. Allele loss with a microsatellite marker locus (NEFL) at 8p12-22 in the five informative tumours from BOV3 suggests a tumour suppressor gene in this region may also be implicated in the pathway of grade 3 breast carcinomas. A high level of allele loss (67 %) at this region was also demonstrated from the analysis of the sporadic breast tumour/lymphocyte DNA pairs (previously assessed for allele loss at BRCA1 and BRCA2 in sections 5.3 and 5.4, respectively: V. Murray, personal communication). Other allele loss studies have suggested the presence of a tumour suppressor gene deleted in breast tumours in

Table 5.3. Similar patterns of allele loss in grade 3 breast tumours from BOV3

Chromosomal regions assessed				
ID	BRCA1	BRCA2	8p12-22	10q24-25
3563	LOH	LOH	LOH	RET
3566	LOH	LOH	LOH	RET
3548	LOH	LOH	LOH	RET
4352	LOH	ND	NI	NI
3677	LOH	LOH	LOH	RET
4351	LOH	ND	ND	RET
3684	LOH	LOH	NI	RET

KEY:

LOH= loss of heterozygosity (allele loss)

RET= retained heterozygosity

NI= constitutively homozygous

ND= not determinable (equivocal)

(The microsatellite marker locus analysed for each region are described in section 5.5.)

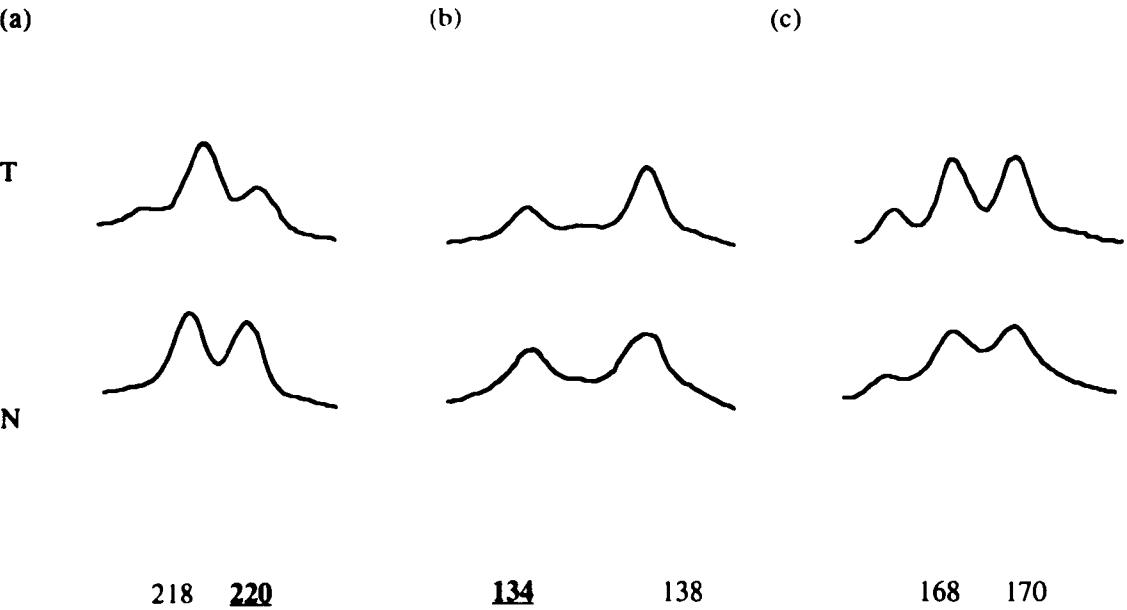
Figure 5.11. Pattern of allele loss in tumour 3548 from the BOV3 family

Examples of the results from microsatellite marker loci analysis of tumour DNA of individual 3548 from the BOV3 pedigree (further details can be interpreted from Figure 3.4). Allele loss of allele 146 at D17S855 (intragenic to BRCA1) in tumour DNA from 3548 is shown in Figure 5.2 (h). The results are comparable to those obtained for breast tumours from other members of the BOV3 pedigree (see Table 5.3). The top allele profile at each locus is from the tumour DNA (T), and the allele profile is that obtained from the lymphocyte DNA (N). Alleles were detected (and sized) as described in section 2.7.3. and allele status interpreted as described in section 2.8. Allele typing (in bp) is shown below each lane; allele in bold and underlined indicates the allele deleted in each tumour (except in (c) where no allele loss as observed).

(a) Loss of allele 224 at SLS234 (maps proximal to BRCA2) in tumour DNA

(b) Loss of allele 134 at NEFL (8p12-22) in tumour DNA

(c) Retention of both alleles at AFM337xf9 (10q24-25) in tumour DNA.



this region (Kerangueven *et al* 1995). Conversely, no allele loss was detected with AFM337xf9 (which maps in the region 10q24-25 which is frequently deleted in prostate tumours: Gray *et al*, 1995) in the six BOV3 breast tumour DNA samples which were informative at this microsatellite marker locus. 30 % allele loss at 10q24-25 was observed from the analysis of the panel of sporadic breast tumours which were previously assessed for BRCA1, BRCA2 and 8p12-22 allele loss (V. Murray, personal communication). An example of the pattern of allele loss detected in the BOV3 breast tumour DNA samples is shown in Figure 5.11.

To evaluate the finding of combined loss of BRCA1/ BRCA2 in the grade 3 ductal breast carcinomas from BOV3, the 118 sporadic paired tumour and lymphocyte DNA samples, previously assessed for BRCA1 and BRCA2 allele loss (sections 5.3 and 5.4), were examined. Allele loss had been assessed without prior knowledge of the histology or other clinical parameter of each carcinoma. A similar pattern of allele loss at BRCA1 and BRCA2 with that observed from the BOV3 tumours was revealed when 45 grade 3 infiltrating ductal carcinomas from the sample set were analysed separately (Table 5.4).

Table 5.4. Distribution of BRCA1 and BRCA2 allele loss among 45 sporadic grade 3 infiltrating ductal breast carcinomas

		BRCA2 allele loss	
		No	Yes
BRCA1 allele loss	No	12	4
	Yes	8	21

A test of independence of allele loss showed strong evidence against this hypothesis (Fisher’s exact test, $p = 0.004$: Dr. P. Smith, personal communication) with an excess of carcinomas being concordant for BRCA1 and BRCA2 loss or retention in grade 3 breast carcinomas. Similar tests of the sporadic carcinomas for grade 1 and grade 2 infiltrating ductal carcinomas or infiltrating lobular carcinomas showed no such trend (seventy three carcinomas in total, data not shown). Also, joint allele loss at BRCA1 and BRCA2 did not correlate with size of tumour at presentation, stage or ER status (data not shown).

These observations support a combined role for BRCA1 and BRCA2 in the tumourigenic pathway of some familial and sporadic breast cancer, in particular, those histologically described as grade 3 (Kelsell *et al*, 1996a). Although a definitive role for BRCA1 and BRCA2 in sporadic tumorigenesis has not been forthcoming, this analysis does suggest that grade 3 breast carcinomas with concordant allele loss at BRCA1 and BRCA2 should be further investigated for germline mutations and/or somatic mutations in these genes. Alternatively, hemizygous loss of BRCA1 and

BRCA2 may be sufficient for the development of sporadic breast tumours. This is discussed further in Chapter 7.

5.7 Summary

Allele loss studies were performed and demonstrated that in breast carcinomas from BRCA1 gene carriers, loss of the wildtype allele is a common event. This supports BRCA1 acting as a tumour suppressor gene. Additionally, a number of ICRF breast cancer pedigrees were assessed for linkage to BRCA1, using a combination of haplotype and allele loss studies. The incorporation of allele loss data increased the power of linkage analysis, particularly for small pedigrees or those with few surviving cases.

Hemizygous loss of BRCA1 and BRCA2 are common events in sporadic breast carcinomas, but, at present, there is no supporting evidence that the retained copy of each gene is somatically mutated. However, analysis of familial and sporadic infiltrating ductal grade 3 breast carcinomas revealed a pattern of combined loss or retention of BRCA1 and BRCA2. This supports a role for both genes in the development of this tumour type.

The observed pattern of LOH at BRCA1 and BRCA2 could be explained by either of the following hypotheses:

- (i) There being at least two types of grade 3 tumours, one of which involves preferential LOH at BRCA1 and BRCA2, or
- (ii) There being at least two types of grade 3 tumours, one of which has a high level of genomic instability indicated by frequent LOH at a number of loci.

This latter hypothesis is independent of BRCA1 and BRCA2 and so can be tested by examining LOH at other chromosomal regions. Allele loss performed on the 45 sporadic grade 3 breast tumour/lymphocyte pairs with microsatellite marker loci mapping to 8p12-22 and 10q24-25 revealed no correlation with combined allele loss/retention at BRCA1 and BRCA2 (data not shown). This supports the former hypothesis and would indicate that BRCA1 and BRCA2 function on compensatory pathways, rather than being involved in the same tumourigenic pathway.

6.1 Concept

An alternative approach to identify cancer susceptibility gene loci is described in this chapter. This approach is based on the observation that, in some families, cancer susceptibility appears to segregate with a distinct autosomal dominant disease phenotype. Linkage studies are performed to localise the fully penetrant disease locus. From identification of this disease locus, it is possible to explore if the same gene defect confers the cancer susceptibility segregating through the family. An alternative explanation is that the two phenotypes may be due to the segregation of two independent but tightly linked gene loci. The latter scenario may be apparent if there are observed cross-overs between the two diseases. This approach has the following advantages over that of mapping a highly but only partially penetrant cancer susceptibility gene:

- (i) the disease manifests usually before puberty, so that gene carriers can be phenotypically identified at an early age
- (ii) the disease susceptibility allele is fully, or almost fully, penetrant
- (iii) the incidence of disease in the population is relatively uncommon compared to, for example, breast cancer, so the problem of phenocopies for the non-malignant phenotype is reduced.

6.2. Keratin diseases

Keratins have been implicated, by morphological and genetic studies, as candidate genes in diseases displaying epidermal fragility (reviewed Lane, 1994). The keratin intermediate filaments form the cytoskeleton of all epithelial cells. Each keratin filament consists of a long α -helical rod domain between non-helical ends. Within the rod domain are a series of hydrophobic amino acid heptad repeats, which allow obligatory heterodimerisation between a type I (acidic) and a type II (basic) keratin forming a

coiled-coil dimer (reviewed Steinhert, 1993). It is these heteropolymers which assemble into the cytoskeletal filaments. Specific type I/type II keratin pairs are co-expressed at particular stages of epithelial differentiation and in a cell specific manner (reviewed Fuchs and Coulombe, 1992).

Keratin mutations have been identified in a number of disorders associated with keratinocyte fragility, such as keratins (KRT) 5 and 14 in epidermolysis bullosa simplex (EBS: Coulombe *et al*, 1991; Lane *et al*, 1992) and KRTs 1 and 10 in bullous ichthyosiform erthroderma (BCIE: McLean *et al*, 1994). In those identified autosomal dominant keratin disorders, the majority of mutations have been found to be point mutations which tend to be clustered in the six domains critical for stable keratin filament heterodimerisation, particularly the 1A rod domain (reviewed Lane, 1994). These keratin mutations are likely to reduce the resistance of the cytoskeleton to minor external trauma leading to collapse of the filament network in epidermal keratinocytes expressing those mutant keratins.

6.3. Palmoplantar keratoderma

The palmoplantar keratodermas (PPK) are a heterogeneous group of skin diseases characterised by hyperkeratosis (“thickening”) of the skin of the palms and soles (reviewed Lucker *et al*, 1994; Stevens *et al*, 1996a). Autosomal dominant forms of the disease have been subclassified according to the presence or absence of epidermolysis (“cell bursting”), the pattern of lesions within the palm and sole skin, associated histopathologies and, more recently, from the identification of the specific gene defect. Broadly, the PPKs fall into three classes based purely on the pattern of distribution of palmoplantar hyperkeratosis: diffuse, focal and punctate (Figure 6.1).

There is only one identified keratin whose expression is specifically restricted to the palmoplantar epidermis, the type I keratin, KRT9 (Langbein *et al*, 1993). Mutations in the 1A region of the KRT9 rod domain have been identified in families with an epidermolytic form of diffuse palmoplantar keratoderma (EPPK; Reis *et al*, 1994). However, of relevance to this study, a large pedigree has been described in which

EPPK occurs in association with an increased susceptibility to breast and/or ovarian cancer (Blanchet-Bardon *et al*, 1987). A KRT9 mutation has been identified which segregates with EPPK (and breast/ovarian cancer) in this family (Torchard *et al*, 1994). The majority of type I keratins, including, KRT9, are clustered on chromosome 17q12-21 and map approximately 1 cM proximal to BRCA1 (Albertsen *et al*, 1995). Thus in this family, it was likely that the KRT9 mutation was segregating with a mutation in BRCA1. Subsequently, a nonsense mutation in exon 11 of BRCA1 has been identified in affected individuals from this family (Friedman *et al*, 1995).

Non-epidermolytic forms of PPK (NEPPK) have been shown to be associated with malignancy. An extensive family from Liverpool, UK has been described in which focal NEPPK (previously termed “Tylosis”) is associated with susceptibility to oesophageal carcinoma (Howell-Evans *et al*, 1958; Ellis *et al*, 1994). Families with punctate NEPPK have also been identified in which the disease is associated with internal neoplasia (Bennion and Patterson, 1984; Bianchi *et al*, 1994; Stevens *et al*, 1996b). Genetic studies in these families may determine if the cancer susceptibility is due to mutations in the PPK disease gene(s) or in a closely linked locus.

6.4. Aim of study

Linkage analyses were performed on families with three autosomal dominant forms of non-epidermolytic palmoplantar keratoderma: diffuse, punctate and focal. For two of these forms, families were identified in which the NEPPK was associated with cancer susceptibility. The aim of the study was to determine the genetic basis of NEPPK predisposition and to assess the relationship between NEPPK and cancer susceptibility.

As all the identified epithelial keratins have been localised to either of two keratin gene clusters: the acidic keratin cluster (type I) at 17q12-21 and the basic keratin gene cluster (type II) at 12q11-13 (Lessin *et al*, 1988; Rosenberg *et al*, 1988), initial linkage studies were performed using microsatellite marker loci flanking these two keratin clusters (detected as described in section 2.7). Microsatellite marker loci were

identified from published genetic maps (Gyapay *et al*, 1994; Dib *et al*, 1996) unless otherwise stated.

RESULTS

6.5. Diffuse NEPPK

Diffuse NEPPK presents with an even, thick, yellow hyperkeratosis over the whole of the palm and sole (Figure 6.1(b)). The disease manifests within the first year of life and is 100 % penetrant. DNA from informative individuals from two families with diffuse NEPPK (designated 2242 and 2244) were analysed with microsatellite marker loci spanning the two gene keratin clusters. No association between NEPPK and malignancy was observed in these pedigrees.

Linkage analysis was performed in family 2244 (as described in section 2.13.2) and the computed LOD scores are shown in Table 6.1. Linkage was observed between the disease and microsatellite marker loci localised to chromosome 12q11-13 with a maximum two-point LOD score of 3.11 at a recombination fraction of 0.00 for D12S368.

Table 6.1. Linkage of diffuse NEPPK to the Type II keratin gene cluster at 12q11-13									
Pedigree	Locus	Recombination Fraction							Candidate genes
		0.0	0.01	0.05	0.1	0.2	0.3	0.4	
2244	KRT9	-8.94	-6.61	-4.11	-2.64	-1.24	-0.56	-0.18	Type I keratins
2244	D12S85	2.53	2.47	2.24	1.93	1.31	0.70	0.19	Type II keratins
2244	D12S368	3.11	3.06	2.87	2.61	2.07	1.46	0.78	

Haplotypes for microsatellite marker loci mapping to chromosome 12q11-13 were constructed for this pedigree and for pedigree 2242. For meioses where phase was equivocal, haplotypes showing the minimum number of recombination events were formed. In both pedigrees studied, there was a haplotype that segregated with the disease (Figure 6.2 (a) and (b)).

Figure 6.1. Three patterns of non-epidermolytic palmoplantar keratoderma

The soles of a representative affected individual from families with either of the three forms of non-epidermolytic palmoplantar keratoderma (NEPPK) analysed in this study.

(a) Focal NEPPK

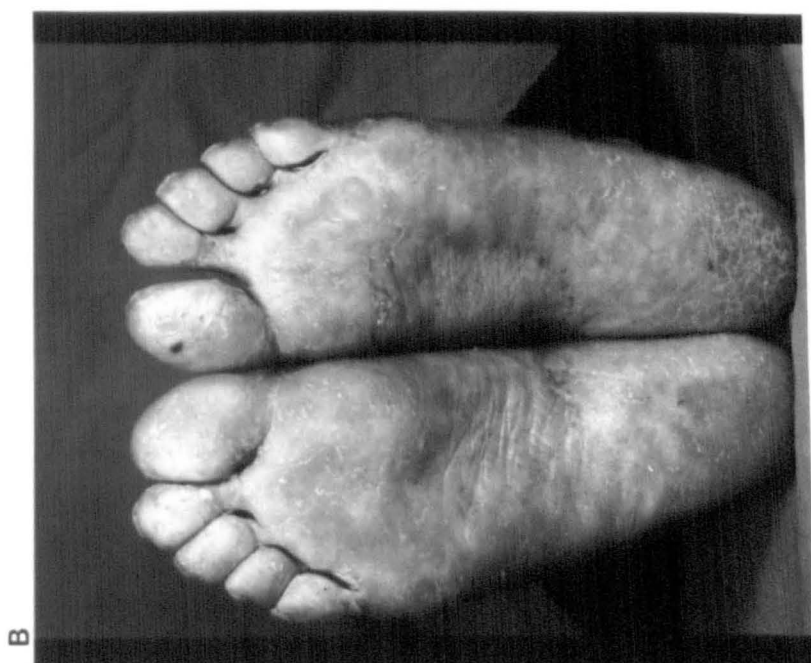
Affected individual from family 2239 has large hard compact masses of keratin at pressure points and other sites of recurrent friction on the sole, in particular, the toes, ball and heel of the foot.

(b) Diffuse NEPPK

Affected individual from family 2242 has an even, thick hyperkeratosis over the whole of the sole.

(c) Punctate NEPPK

Affected individual from family 2039 has small punctate keratoses over the whole of the sole.



As germline mutations in the type I keratin, KRT9, have been demonstrated in a number of diffuse epidermolytic PPK pedigrees (Reis *et al*, 1994; Torchard *et al*, 1994), it has been proposed that all patterns of diffuse PPK are the result of KRT9 mutations. However, this study shows that diffuse NEPPK is linked to the region 12q11-q13 where the type II keratins are located. As KRT9 maps to 17q12-21, a mutation in KRT9 can be excluded as the inherited predisposition to diffuse NEPPK. This confirms genetically that diffuse NEPPK and EPPK are separate disease entities. Linkage to the same region on chromosome 12 with diffuse NEPPK was also observed in families from northern Sweden (Lind *et al*, 1994).

A proposed candidate gene for this disease is the type II keratin, KRT1 (Kimonis *et al*, 1994), which maps centromeric to the microsatellite marker locus D12S96 and telomeric to D12S803 (Figure 6.2(c)). No KRT1 DNA sequence variants have been detected in DNA from affected family members when compared to that obtained from unaffected individuals (Prof. I. Leigh, personal communication). A crossover in one affected individual from family 2244 places the disease locus centromeric to the microsatellite marker locus D12S803 (arrowed in Figure 6.2(a)). This recombination event excludes a number of keratins, including KRT1, as candidate genes for diffuse NEPPK on the basis of their physical location (Figure 6.2(c)). KRT 7 could be regarded as a potential candidate gene as it maps proximal to D12S803 (Figure 6.2(c)). However, this keratin is biologically unlikely to be the genetic basis of this disease as it is expressed in, for example, the mammary gland and bladder epithelial cells but not in the palmoplantar epidermis (Moll *et al*, 1988).

It is plausible that a mutation in a KRT 6 gene may be the genetic basis of NEPPK as it may map proximal to D12S803 (Figure 6.2(c)), but its expression appears not to be limited to palmoplantar skin (Moll *et al*, 1988). However, five KRT 6 sequences have been identified (two functional and three likely to be pseudogenes: Prof. I. Leigh and Dr. I. McLean, personal communications) and it is possible that one of these KRT6 genes is expressed predominantly in the palmoplantar epidermis. The position of KRT6 genes in relation to D12S803 and their DNA sequence analysis in diffuse

Figure 6.2. Mapping of diffuse NEPPK to the Type II keratin gene cluster

Haplotype analysis of two pedigrees in which diffuse NEPPK is segregating with microsatellite marker loci spanning the type II epithelial keratin gene cluster on chromosome 12q11-13. (The allele numbering is not comparable between the two families).

(a) Structure and haplotype analysis of pedigree 2244.

The sex of each individual is as described in Figure 1.1. Affected status is indicated by a red-filled in box/circle. The haplotypes are shown under each individual. The microsatellite marker loci genotyped and their order are shown beneath the pedigree. A zero allele indicates individual not genotyped at that microsatellite marker locus. The haplotype segregating with the disease in each family is indicated in red text. The affected individual arrowed has inherited the disease haplotype centromeric to D12S803.

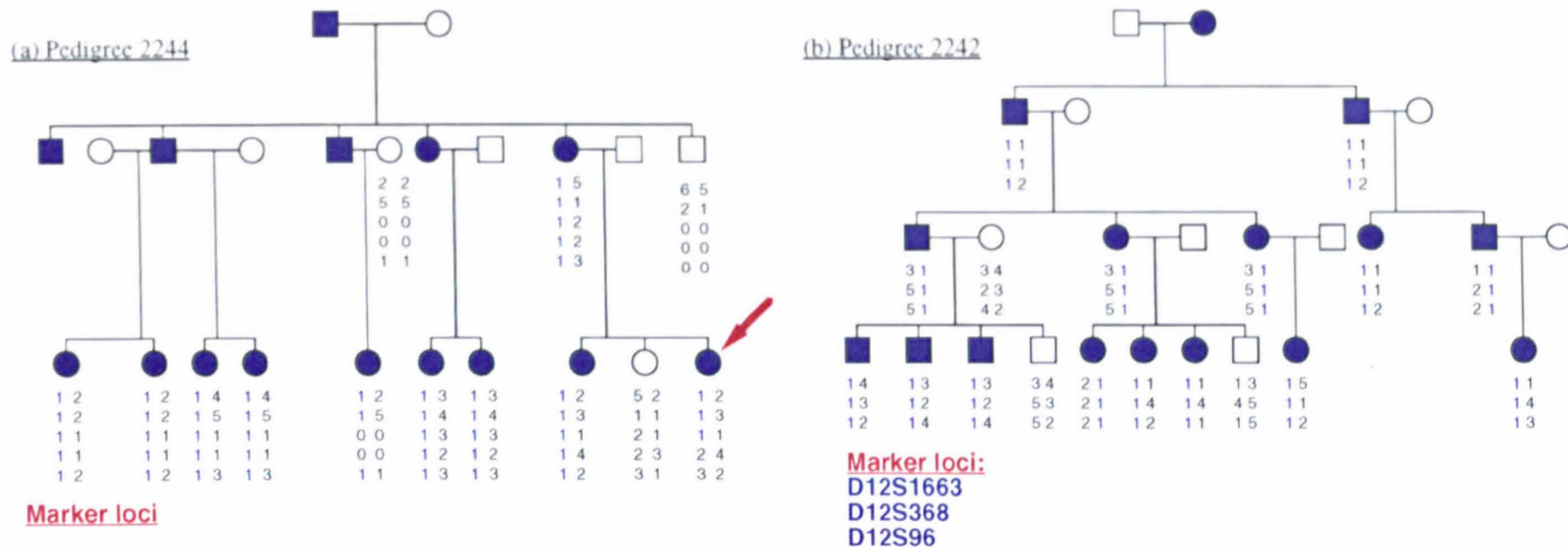
(b) Structure and haplotype analysis of pedigree 2242

The pedigree is constructed and labelled as in Figure 6.2(a). The haplotype segregating with the disease is indicated in red text.

(c) Organisation of the type II keratin cluster

Schematic representation of the organisation of the keratin genes in relation to microsatellite marker loci mapping to chromosome 12q11-13 (adapted from Krauter *et al*, 1995). Microsatellite marker loci are indicated above the bold line with the position of each keratin gene below the line. The orientation of the loci in relation to the position of the centromere (CEN) and telomere (TEL) is indicated. The arrow above the line indicates the diffuse NEPPK disease gene maps centromeric to D12S803 (from the cross-over detected in the affected individual arrowed in pedigree 2244 (Figure 6.2(a)).

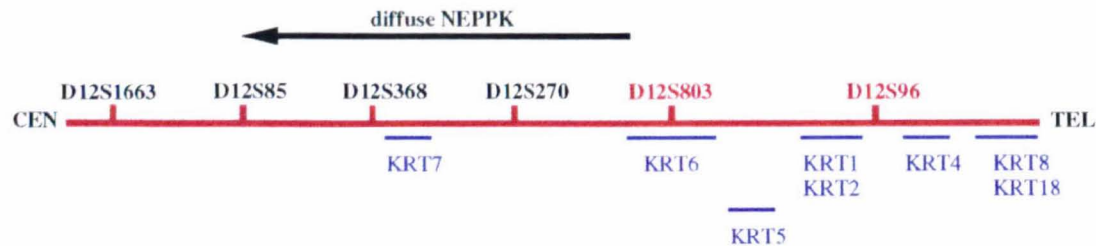
Figure 6.2 Mapping of diffuse NEPPK to the Type II keratin gene cluster



Marker loci

D12S85
D12S368
D12S270
D12S803
D12S96

(c) Organisation of the type II keratin cluster (adapted from Krauter *et al.*, 1995)



KEY:

- diffuse NEPPK
□ unaffected

NEPPK affecteds may identify such a candidate. Alternatively, a mutation in a unidentified keratin (or a gene which effects keratin integrity) may underly diffuse NEPPK. Further studies are in progress to identify the genetic basis of this disease.

6.6 Focal NEPPK without associated Oesophageal Cancer

Focal NEPPK presents as large hard compact masses of keratin particularly on the sole and also at other sites of recurrent friction (Figure 6.1(a)). The NEPPK in the pedigree studied, 2239 (see Figure 6.3), was inherited as an autosomal dominant trait and manifested at around 7 years of age. Intermittent subtle nail abnormalities and oral hyperkeratosis were also observed in affected individuals from the pedigree (Dr. H. Stevens, personal communication). Linkage analysis was performed as in section 2.13.2.

Linkage and haplotype analysis of the focal NEPPK pedigree provided evidence for the localisation of the disease gene to chromosome 17q12-21, the location of the type I keratin gene cluster (Table 6.2; Figure 6.3.(a)). A two-point LOD score of 2.49 was obtained with the disease and a microsatellite marker mapping within intron 4 of the KRT9 gene (Reis *et al*, 1994) at a recombination fraction of 0.00. A three-point analysis (as described in section 2.13.2) with KRT9, D17S855 and the disease gave a LOD score of 3.25 at a recombination fraction of 0.0 (Table 6.2), providing stronger evidence for linkage of focal NEPPK to this region. There was clear evidence against linkage of the disease to the type II keratin gene cluster at 12q11-13, with a negative LOD score of -2.22 with the marker locus D12S85 at a recombination fraction of 0.1 (Table 6.2; Figure 6.3 (b)).

Table 6.2. Linkage of focal NEPPK to the Type I keratin gene cluster at 17q12-21

Pedigree	Locus	Recombination Fraction							Candidate genes
		0.0	0.01	0.05	0.1	0.2	0.3	0.4	
2239	KRT9"CA"	2.49	2.44	2.27	2.04	1.56	1.06	0.54	Type I keratins
2239	D17S855	2.25	2.20	1.98	1.71	1.17	0.65	0.20	
2239	KRT9/D17S855	3.25	3.20	2.97	2.68	2.07	1.42	0.73	
2239	D12S85	-8.37	-5.68	-3.39	-2.22	-1.09	-0.50	-0.17	Type II keratins
2239	D12S368	-5.81	-2.76	-1.43	-0.90	-0.43	-0.20	-0.07	

A disease-causing mutation in the type I keratin KRT9 (genetic basis of EPPK: see section 6.3) is unlikely in this family, as this keratin is only expressed in palmoplantar skin and therefore would not explain the distribution of oral lesions plus the other observed secondary characteristics. The expression of KRT's 6 and 16 in the palmoplantar epidermis and oral mucosa (Moll *et al*, 1988) strongly correlates with the clinical lesions present in affected individuals from this pedigree. However as KRT6 is a type II keratin mapping to 12q11-13, this implicates KRT16, which maps to 17q12-21, as a strong candidate gene for focal NEPPK. Subsequent DNA sequence analysis revealed mutations in the 1A domain of KRT16 in affected members of this and another focal NEPPK family (Shamsher *et al*, 1995). This domain is critical for keratin filament formation and integrity (McLean and Lane, 1995).

Mutations in the 1A domain of KRT16 (and KRT17) have also been demonstrated to occur in the clinically similar disease pachyonychia congenita (McLean *et al*, 1995). A significant clinical difference between the two skin diseases is the gross thickening and curvature of the nails seen in pachyonychia congenita affected individuals which is absent in focal NEPPK. This may reflect a functional difference in the amino acids substituted in KRT16 in each disease (Shamsher *et al*, 1995). In addition to the site of mutation, severity of phenotype may be related to distinct "modifying" genetic factors or due to environmental effects such as physical trauma.

Figure 6.3. Mapping of focal NEPPK (without associated Oesophageal Cancer) to the Type I keratin gene cluster

Haplotype analysis of pedigree 2239 with microsatellite marker loci spanning the two epithelial keratin gene clusters. The segregation of a 17q12-21 haplotype with focal NEPPK supports the localisation of the disease gene in this family to the chromosomal region harbouring the type I keratin genes.

(a) Structure and haplotype analysis of 17q12-21 in pedigree 2239.

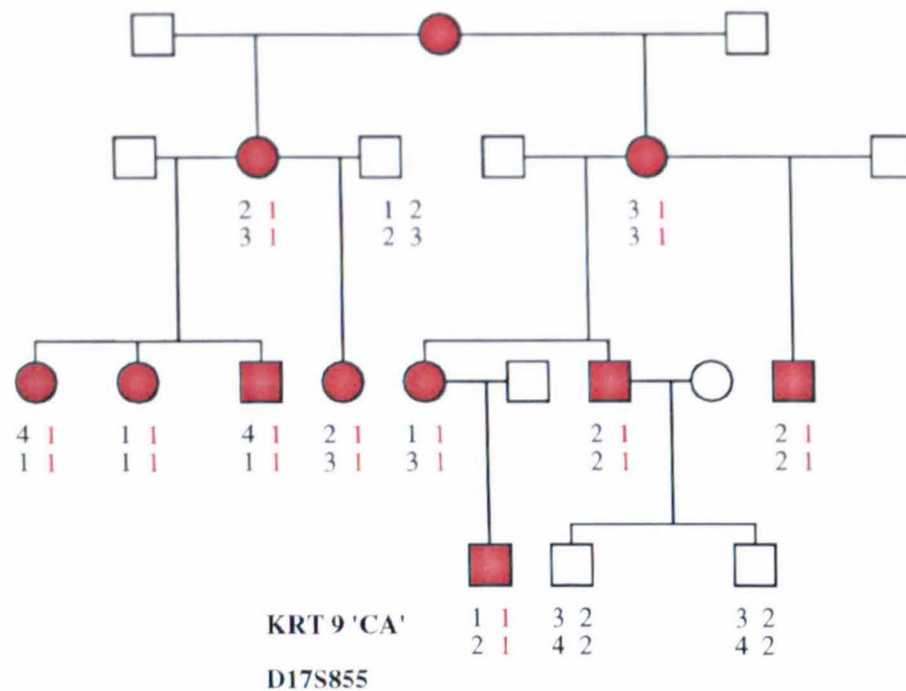
The sex of each individual is as described in Figure 1.1. Affected status is indicated by a red-filled in box/circle. The haplotypes are shown under each individual. The microsatellite marker loci genotyped and their order are shown to the left of the affected typed individual in generation 4. The microsatellite marker locus, designated KRT9 'CA' maps within intron 4 of the KRT9 gene (Reis *et al*, 1994). The haplotype which is segregating with the disease is indicated in red text.

(b) Structure and haplotype analysis of 12q11-13 in pedigree 2239.

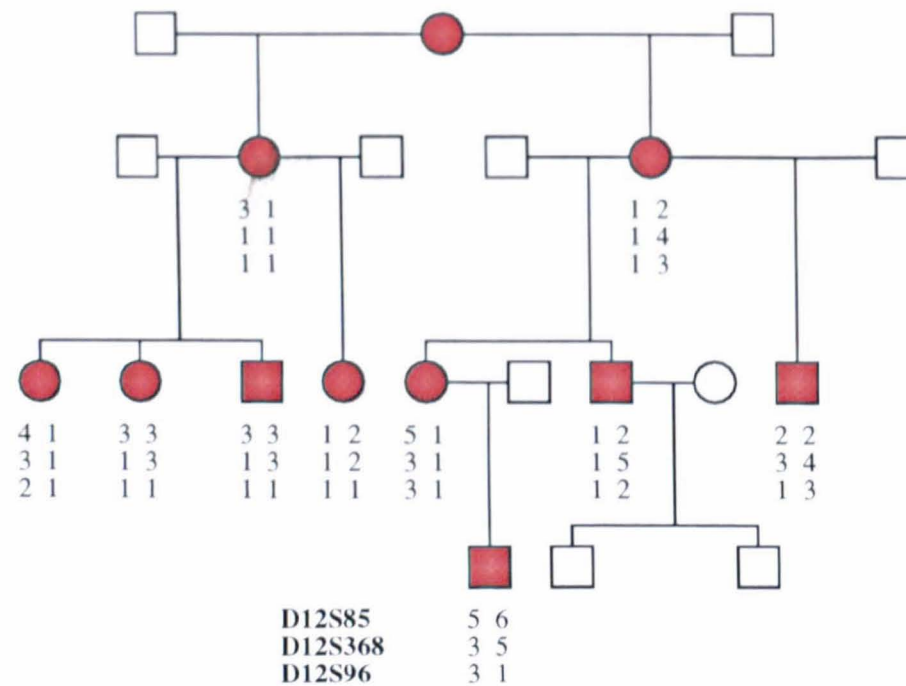
Pedigree is constructed and labelled as in Figure 6.3(a). The physical mapping of these microsatellite marker loci in relation to the type II keratin genes is shown in Figure 6.2(c).

Figure 6.3. Mapping of Focal NEPPK (without associated Oesophageal Cancer) to the Type I keratin gene cluster

(a) Pedigree 2239: 17q12-21 Type I keratins



(b) Pedigree 2239: 12q11-13 Type II keratins



KEY:



Focal NEPPK



Unaffected

6.7. Focal NEPPK and Oesophageal Cancer

A large pedigree, designated 2216, from the United States of America (USA) was studied in which focal NEPPK was segregating in association with susceptibility to squamous cell carcinoma (SCC) of the oesophagus. A number of the branches of this family are shown in Figure 6.4. Other malignancies were also present in certain family members presenting with the skin disease. The focal NEPPK disease phenotype segregated as an autosomal dominant trait with 100% penetrance. The pattern of focal NEPPK in this family is clinically similar to that for focal NEPPK affected individuals who carry a mutation in KRT16 (described section 6.6). However, the normal appearance of the nails in affected individuals from this family is one noticeable clinical difference between the two pedigrees (not shown). The nails from KRT16-focal NEPPK affected individuals have an obvious widening of the onychocorneal band and terminal splinter haemorrhages (Dr. H. Stevens, personal communication). From family 2216, 125 affected individuals in seven generations have been identified with 17 affected individuals having associated cancer (the complete pedigree is described in Stevens *et al*, 1996a). Eight cases of SCC of the oesophagus were identified in individuals with focal NEPPK, of which seven occurred in heavy tobacco smokers or chewers. An NEPPK affected individual from this family has a 40% chance of developing oesophageal cancer by age 70 (Prof. T. Bishop, personal communication). Other malignancies in NEPPK affected individuals from this family included two cases of malignant melanoma, two cases of carcinoma of the breast, two cases of lung carcinoma, two cases of colonic adenocarcinoma, one case of chronic lymphocytic leukaemia and one case of multiple adenomatous polyps of the colon; none of which sites were significantly increased in frequency (Prof. T. Bishop, personal communication).

6.8 Linkage analysis in pedigree 2216

Blood samples were drawn and DNA extracted from affected and unaffected members of the focal NEPPK and oesophageal cancer family 2216 (ascertained and extracted as described in sections 2.1.2 and 2.3.2, respectively). DNA samples were genotyped by






PCR amplification with microsatellite marker loci flanking the two keratin clusters at 12q11-13 and 17q12-21 (same loci as described in sections 6.5 and 6.6). Alleles were detected as described in section 2.7.3. A number of cross-over events between NEPPK and these microsatellite marker loci in affected (and unaffected) individuals indicated that disease in this family was not due to an inherited mutation in the candidate genes KRT6 or 16 (or other genes) mapping in these two chromosomal regions (data not shown).

A large extensive family from Liverpool, UK in which “Tylosis” (a form of NEPPK) is associated with a high risk of developing oesophageal cancer has been described (Howell-Evans *et al*, 1958). It has been calculated that a NEPPK affected individual in this family has a 92% probability of dying from oesophageal cancer by age 70 (Ellis *et al*, 1994). In this family, genetic linkage was demonstrated with the disease and microsatellite marker loci mapping to 17q23-qter locus (Risk *et al*, 1994). This was termed the Tylosis Oesophageal Cancer (TOC) locus. As this region maps approximately 45 cM telomeric to the keratin gene cluster on chromosome 17 (Dib *et al*, 1996), a mutation in the type I keratins including KRT16 can be genetically excluded as the inherited basis of this pattern of keratoderma. Interestingly, a clinical reappraisal of the “tylotic” phenotype in affected members of this family indicated that it is best described as focal NEPPK similar to that seen in the focal NEPPK/oesophageal cancer pedigree 2216 (Prof. I. Leigh, personal communication).

Lymphocyte DNA from 2216 family members was genotyped with microsatellite marker loci mapping to the 6-7 cM minimal region for the TOC locus defined from haplotype analysis of the UK pedigree (flanked by D17S929 and D17S937, Risk *et al*, 1994: Figures 6.4 and 6.5). Microsatellite marker loci were identified from published genetic maps (Gyapay *et al*, 1994; Dib *et al*, 1996) and linkage analysis was performed as described in section 2.13.2. A maximum two-point LOD score of 8.20 at a recombination fraction of 0.00 for the microsatellite marker locus D17S1603 and focal NEPPK was calculated. LOD scores obtained for D17S1603 and D17S785 are shown

Figure 6.4. Mapping of focal NEPPK/oesophageal cancer in pedigree 2216 to the TOC region

Haplotype analysis of pedigree 2216 with microsatellite marker loci mapping to the region harbouring the TOC locus. The pedigree is as described in the box in the left hand corner of the first page of the pedigree. The age of diagnosis of malignancy (if known) is shown to the right of the cancer type under (or to the right of) the affected individual. Microsatellite marker loci genotyped and their genetic map order is indicated to the left of the figure under the Key. Genotypes at these loci are shown under each typed individual (in the order defined at left of the figure). A zero allele indicates individual not genotyped at that microsatellite marker locus. The putative disease haplotype segregating through the pedigree is indicated in red text.

 Male  NEPPK
 Female  NEPPK and MALIGNANCY:
 Deceased

oe	oesophageal cancer
br	breast cancer
mm	malignant melanoma
co	colon cancer
mo	mouth carcinoma
ut	uterine cancer

Marker order:

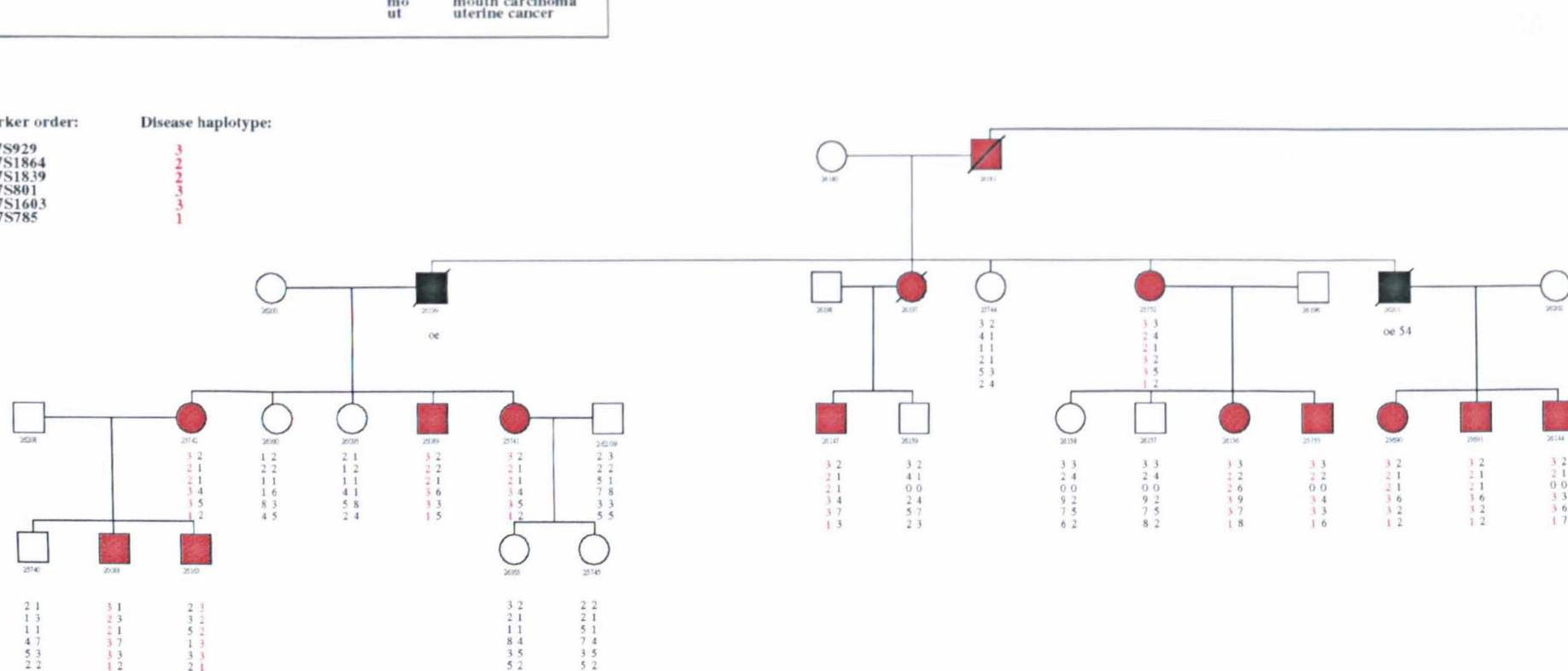
Disease haplotype:

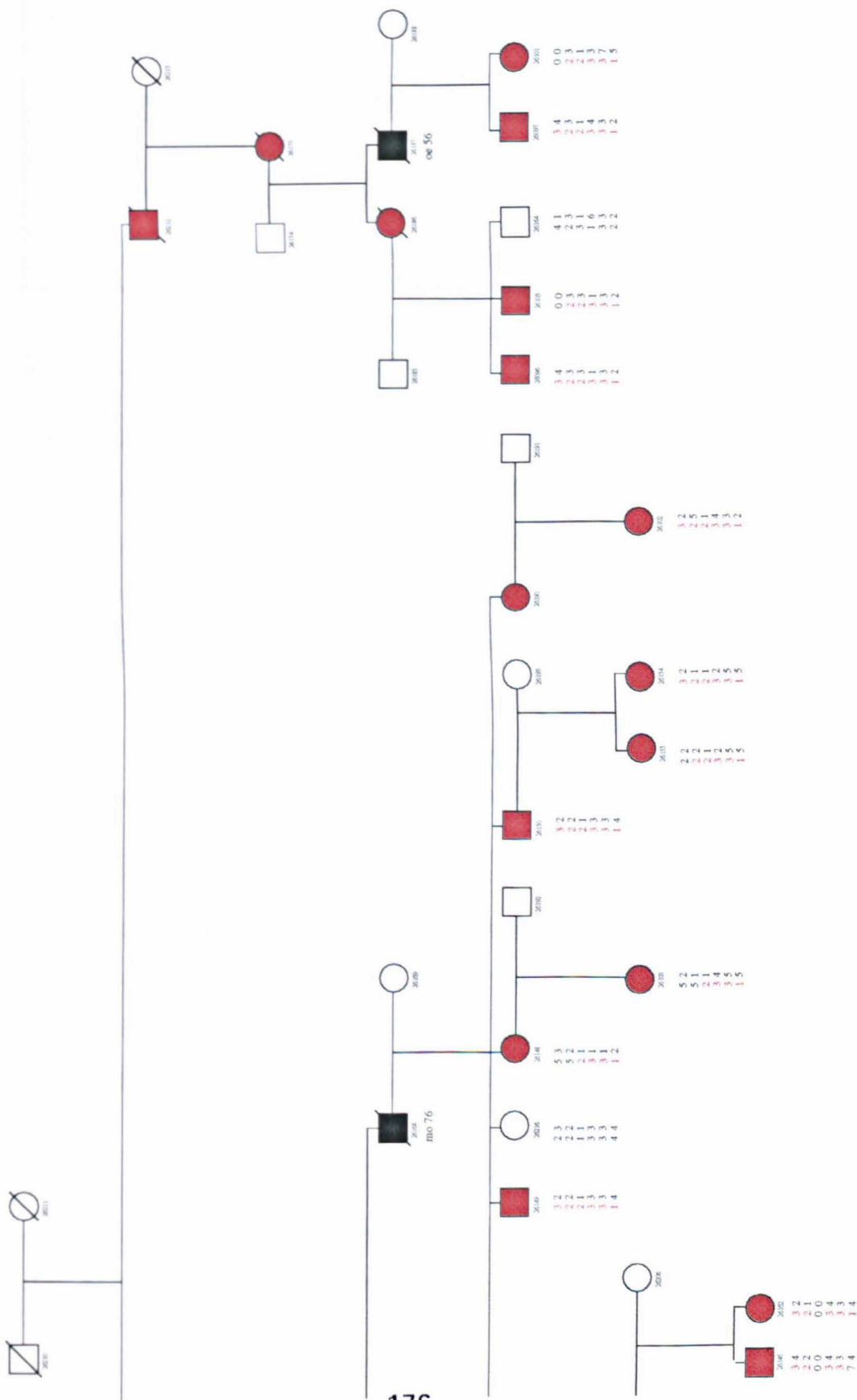
D17S929
D17S1864
D17S1839
D17S801
D17S1603
D17S785

3
2
2
3
3
1

174

Figure 6.4. Mapping of focal NEPPK/oesophageal cancer in Pedigree 2216 to the TOC region





in Table 6.3. This result localises the disease gene segregating in family 2216 to the same region as that of the UK focal NEPPK (“Tylosis”) pedigree (Stevens *et al*, 1996a). This suggests that mutations in the TOC gene underlie the genetic basis of focal NEPPK (and possibly oesophageal cancer predisposition) in these two pedigrees

Table 6.3. Linkage analysis of focal NEPPK with microsatellite marker loci mapping in the region harbouring the TOC locus in pedigree 2216.

Locus	Recombination Fraction							
	0.0	0.001	0.01	0.05	0.1	0.2	0.3	0.4
D17S1603	8.20	8.19	8.02	7.29	6.36	4.51	2.75	1.21
D17S785	6.67	6.65	6.51	5.89	5.11	3.56	2.11	0.87

6.9 Fine mapping of the TOC locus

Linkage analysis demonstrated that the skin disease in family 2216 is linked to microsatellite marker loci mapping in the TOC region. Towards refining the localisation of the TOC locus, haplotypes for individuals from the pedigree 2216 were constructed for microsatellite marker loci mapping in the previously defined minimal region for the TOC locus (Risk *et al*, 1994). Additional family members not previously genotyped for the linkage study (section 6.8) were also analysed. Haplotypes were compared between affected and unaffected individuals, and, if phase was equivocal, they were constructed on the basis of a minimum number of recombination events at meiosis.

Haplotype analysis of this family (and the UK family: Dr. J. Risk, personal communication) confirmed the order of microsatellite marker loci (Dib *et al*, 1996) with the exception that D17S801 was more likely to map between D17S1839 and D17S1603, rather than telomeric to D17S785. An example showing the position of D17S801 centomeric to D17S785 can be seen from haplotype comparison of the offspring of 26152 and 26206 (Figure 6.4). Individual 26146 has inherited the paternal chromosome (3/4: 2: 0: 3: 3: 7). As 26146 has inherited the “7” allele rather than the “1” allele at D17S785 from her father 26152, this suggests a cross-over

telomeric to D17S1603 and D17S801 had occurred. The position of D17S801 centromeric to D17S1603 was obtained from haplotype analysis of the UK family (Dr. J. Risk, personal communication; Kelsell *et al*, 1996b).

A haplotype in pedigree 2216 was clearly demonstrated to segregate with the disease telomeric to the locus D17S1864 and proximal to the locus D17S785 (Figure 6.4). The genetic distance between these microsatellite marker loci is estimated to be 5 cM (Dib *et al*, 1996). Examples of recombination events in affected and unaffected individuals supporting this refined genetic location for the TOC locus can be seen in Figure 6.4. An ideogram in Figure 6.5 summarises these cross-over events and the localisation of the disease gene in this family.

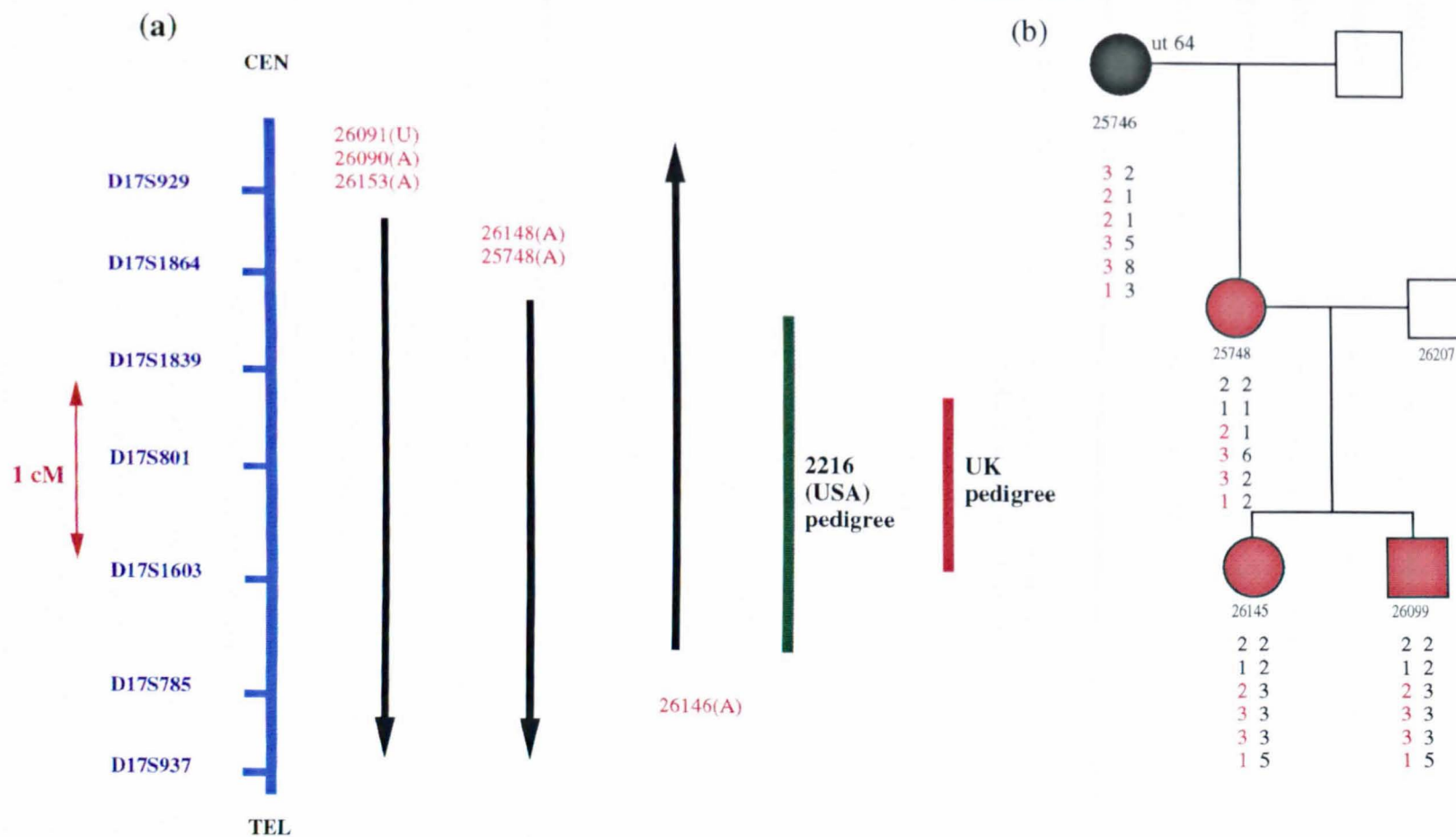
The refinement of the TOC locus from analysis of the USA pedigree 2216 is in agreement with the localisation of the disease locus from haplotype analysis of affected and unaffected members of the UK family (Dr. J. Risk, personal communication; Kelsell *et al*, 1996b). Further recombination events were identified in the UK family which placed the disease gene in a region estimated to be approximately 1 cM, flanked by D17S1839 and D17S1603 (Figure 6.5). Analysis of the UK family was conducted by Dr. J. Risk.

In addition, a third, less extensively characterised pedigree with focal NEPPK (of which two PPK affecteds had developed oesophageal cancer) from Germany has been described. In this pedigree, PPK has also been demonstrated to be linked to microsatellite marker loci mapping to the TOC region (Hennies *et al*, 1995a). Therefore, three families have been identified in which focal NEPPK/oesophageal cancer is linked to microsatellite marker loci mapping to the chromosomal region 17q24-25; designated the TOC locus.

Figure 6.5. Summary of recombination events at the 17q24-25 TOC locus identified in the American focal NEPPK/oesophageal cancer pedigree 2216.

(a) Diagrammatic summary of the localisation of the TOC locus from haplotype analysis of family 2216 (and the UK pedigree; Dr. J. Risk, personal communication). The order of the microsatellite marker loci are shown to the left in orientation with the centromere (CEN) and the telomere (TEL). The TOC locus was previously mapped to a 6-7 cM region flanked by D17S929 and D17S937 (Risk *et al*, 1994). To the right of the map are arrows indicating the position of the TOC locus from specific recombination events identified in affected (A) and unaffected (U) members of the pedigree 2216 (see Figure 6.4). The individual identification numbers and affected status are shown above the arrows. For example, the cross-over identified in individual 26146 positions the TOC locus centromeric to D17S785 (this individual shares the disease haplotype proximal to this locus; Figure 6.4). From the analysis of pedigree 2216, the TOC locus can be localised between D17S1864 and D17S785, indicated by the green bar to the right of the figure. The localisation of the TOC locus from the UK pedigree analysis (Dr. J. Risk, personal communication) between D17S1839 and D17S1603 is indicated by the red bar. This is estimated to be a genetic distance of 1 cM (Dib *et al*, 1996) and is shown to the right of the genetic map.

Figure 6.5. Summary of recombination events identified at the 17q24-25 TOC locus in the American focal NEPPK/oesophageal cancer pedigree, 2216



6.10. Ancestral TOC mutation?

Disease haplotypes were constructed for five gene carriers from each of the three identified TOC-linked focal NEPPK families (USA, UK and German pedigrees; section 6.9). Haplotypes were compared to determine whether the same disease haplotype was segregating in two or more of these families. The microsatellite marker loci analysed spanned the TOC locus between (and including) D17S1864 and D17S785 (Figure 6.5(a)). The allele sizes for the three disease haplotypes were estimated as described in section 2.7.3 and are shown in Table 6.4.

Table 6.4. Disease haplotype comparison of the three TOC-linked focal NEPPK/oesophageal cancer pedigrees

Marker loci	USA	UK	German
D17S1864	120	118	122
D17S1839	255	247	247
D17S801	278	266	266
D17S1603	229	235	235
D17S785	209	189	191

Allele sizes in base-pairs (bp). Alleles in common indicated in bold text.

Interestingly, the disease haplotypes for the UK and German families are identical for the microsatellite marker loci mapping in the minimal region defined for the TOC locus from haplotype analysis: D17S1839, D17S801 and D17S1603 (Figure 6.5(a)). This would support the possibility that the same ancestral mutation is segregating through the two pedigrees. The alleles shared by these two families at the D17S1839 and D17S801 loci are common (55 % and 30 %, respectively: Genome Database). This is in agreement with the observation of a number of homozygotes and spouses carrying these alleles from analysis of the extensive UK pedigree (Dr. J. Risk, personal communication). However, the sharing of a less frequently occurring allele at the D17S1603 locus (17%: 11/66 chromosomes from a panel of 33 unrelated caucasians

as described in section 2.14; data not shown) supports the two families having inherited the same ancestral mutation. If this is so, it also confirms the localisation of the TOC gene to the minimal region defined from haplotype analysis (flanked by D17S1839 and D17S1603).

The American pedigree (2216) has a different disease haplotype compared to the UK and German families. This suggests affected individuals from 2216 are unlikely to be carrying the same mutation as those from the other two pedigrees. The apparent genetic unrelatedness of the American family with the German family is interesting as the lineage of the American family can be traced back to focal NEPPK affected individuals in Germany (though not to members of the German family described in this study: Dr. H. Stevens, personal communication). No genealogical link has been demonstrated between the UK and German pedigree even though they appear to have inherited the same ancestral mutation.

Two lines of evidence support the hypothesis that the same gene defect confers focal NEPPK and oesophageal cancer susceptibility in these pedigrees (compared to the possibility of two closely linked gene mutations segregating in the pedigrees):

- (i) the two phenotypes segregate together in all three pedigrees studied (two of which are extensive (USA and UK families: 6-7 generations)
- (ii) two distinct haplotypes segregate with the disease gene in the three families studied.

6.11 Does TOC act as a tumour suppressor gene?

The question of why the elevated risk of oesophageal squamous cell carcinoma in individuals carrying a TOC mutation and not those who inherited a KRT16 mutation arises. One explanation could be that the TOC gene may function as a tumour suppressor gene. To explore this possibility, paired oesophageal tumour/normal DNA samples from both the USA and UK TOC-linked pedigrees were genotyped with microsatellite marker loci spanning the TOC locus to assess the rate of allele loss as

described in section 2.8 DNA extraction and genotyping was as described in sections 2.3.3 and 2.7.3. In order to determine if the TOC locus has a role in sporadic as well as familial tumour progression, a panel of sporadic oesophageal tumours were also assessed for allele loss.

(i) TOC-linked oesophageal tumours

The initial focus was the comparison of DNA extracted from archival oesophageal tumour sections with matched lymphocyte DNA of affected family members of the USA and UK families. This would provide the strongest evidence for the TOC gene functioning as a tumour suppressor (in an analogous fashion to BRCA1: see section 5.1).

However, reproducible results were only achieved from the DNA analysis of one familial oesophageal tumour/ blood DNA pair. This DNA was extracted from an oesophageal carcinoma from an affected member of the UK family (individual IV 58; pedigree not shown). No evidence of allele loss was detected in the oesophageal tumour DNA at four constitutively heterozygous microsatellite marker loci spanning the TOC locus: D17S929, D17S1864, D17S801 and D17S785 (see map in Figure 6.5). An example of retention of heterozygosity at the D17S929 locus in this tumour is shown in Figure 6.6(a). This result suggests that TOC may not act as a familial tumour suppressor gene.

(ii) Sporadic oesophageal tumours

From the previous section, there is suggestive evidence that TOC does not function as a familial tumour suppressor gene. To determine whether this may also be the case in sporadic cancer, a panel of sixteen archival oesophageal tumour/ normal oesophageal DNA pairs were examined for allele loss. In addition to analysis of allele loss at the TOC locus (D17S929, D17S1864 and D17S785), the microsatellite marker locus, D17S579, was screened across the paired DNA samples to assess allele loss in the chromosomal region harbouring BRCA1. D17S579 maps approximately 1-2 cM telomeric to BRCA1 (see Figure 3.12; Albertsen *et al*, 1994).

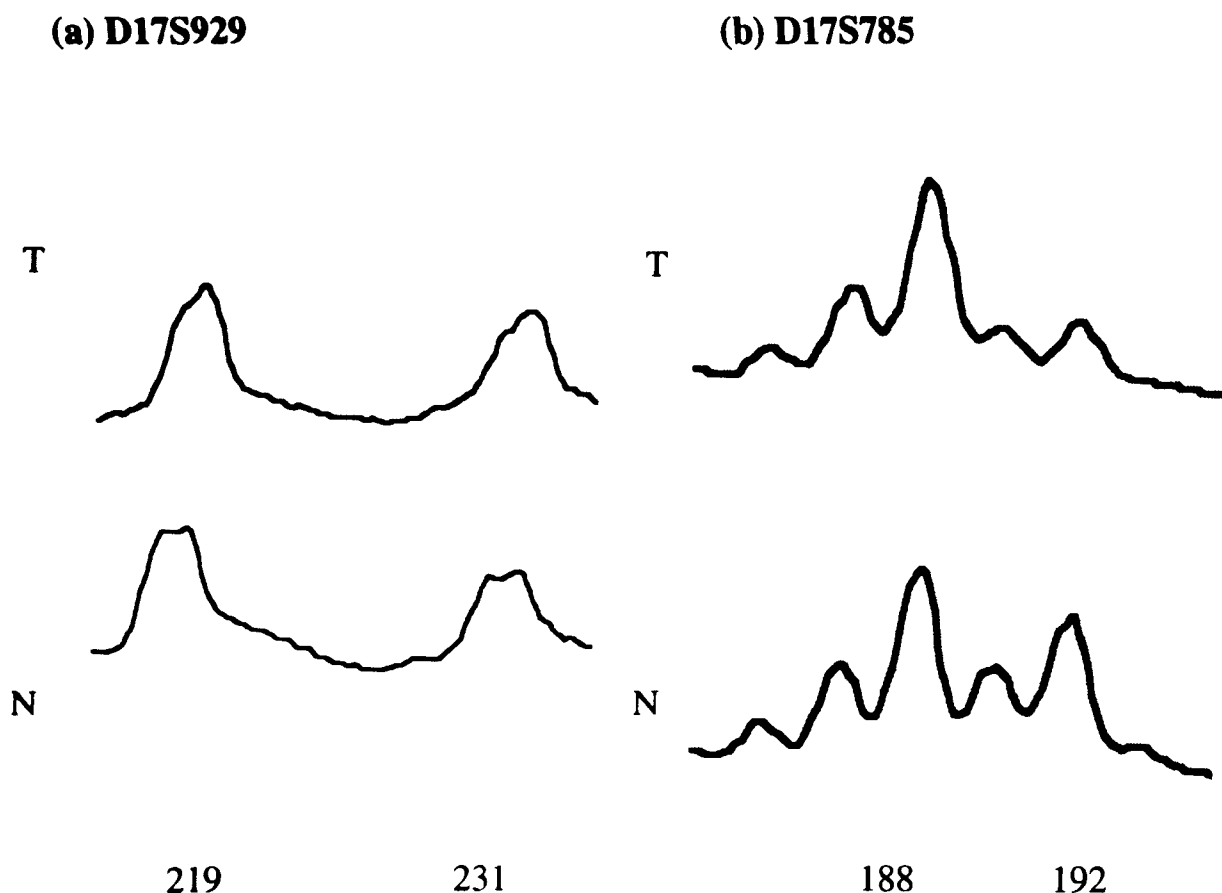


Figure 6.6. Examples of allele status in the chromosomal region 17q24-25 (TOC locus) in oesophageal carcinomas

Examples of the results from microsatellite marker loci analysis of oesophageal tumour DNA. Allele size (in bp) is shown underneath each allele peak (detected and sized as described in section 2.7.3). Allele profiles of PCR amplification products are shown to the right of tumour (T) and normal (N) DNA. Allele status was assessed as described in section 2.8.

(a) No allele loss at D17S929 in a TOC-linked oesophageal tumour

The oesophageal tumour/lymphocyte DNA pair was analysed from an NEPPK affected individual (IV58) who had inherited the TOC haplotype which was segregating with the skin disease (and oesophageal cancer) in the UK family (Dr. J. Risk, personal communication). The allele segregating with disease in the UK family at this locus is the 231 bp allele. Tumour DNA was extracted from archival pathological material (as described in section 2.3.3). No loss was apparent in the tumour DNA when its allele profile was compared to that obtained from the lymphocyte DNA of this individual at this locus.

(b) Allele loss at D17S785 in the oesophageal tumour 8847 (see Table 6.5)

Tumour and normal oesophageal DNA was extracted from archival pathological material (determined as described in section 2.2.5). A reduction of the 192 bp allele peak is clearly evident when compared to that obtained for the normal oesophageal DNA from this individual.

Tumour	Region: BRCA1	17q24-25 (TOC)		
	Marker loci: D17S579	D17S929	D17S1864	D17S785
767	LOH	ND	LOH	ND
8087	NI	LOH	NI	LOH
7769	LOH	NI	RET	NI
4350	LOH	LOH	LOH	NI
20610	NI	LOH	NI	NI
8122	ND	RET	RET	RET
4372	RET	NI	RET	RET
9231	LOH	RET	RET	RET
9641	ND	NI	LOH	LOH
84020	ND	NI	NI	LOH
2039	LOH	NI	LOH	LOH
95770	ND	ND	ND	LOH
4873	LOH	NI	LOH	ND
1693	RET	RET	LOH	ND
8847	NI	NI	NI	LOH
6717	ND	LOH	LOH	NI

KEY:

LOH= loss of heterozygosity (allele loss)

RET= retained heterozygosity

NI= not informative as constitutively homozygous

ND= not determinable

Table 6.5. 17q allele loss in sporadic oesophageal carcinomas

The results are summarised in Table 6.5 with 14 of the 16 oesophageal tumours analysed displaying allele loss at both or one of the chromosomal regions investigated on the long arm of chromosome 17. This high percentage of tumours displaying allele loss (87 % of tumours assessed) does suggest that there is a major locus (or loci) on chromosome 17 implicated in the development of oesophageal carcinomas. An example of allele loss at D17S785 in tumour 8847 is shown in Figure 6.6(b). Three of the oesophageal tumours demonstrated allele loss in only one of the regions tested (Table 6.5). DNA from oesophageal tumours (7769, 9231) demonstrated allele loss specifically at D17S579; whilst tumour 1693 showed loss specifically at D17S1864 which maps telomeric to D17S929. However, the number of tumours studied here is too small to determine which region on chromosome 17 is the focus for allele loss in oesophageal tumours. Mori *et al* (1994) suggested the focal region for allele loss in oesophageal tumours is at a locus mapping adjacent to BRCA1.

In summary, there is no strong evidence for or against TOC acting as a tumour suppressor gene. However the high frequency of 17q allele loss detected in sporadic oesophageal tumours supports the presence of a tumour suppressor gene(s) which has a key role in oesophageal tumorigenesis. Once TOC has been identified, its possible involvement in oesophageal cancer predisposition and/or progression may be determined.

6.12 Candidate genes for TOC?

A number of biologically plausible genes localised to the long arm of chromosome 17 can be excluded as candidates for TOC, including KRT16 and the other type I keratins, as they map approximately 45 cM centromeric to the TOC locus (Dib *et al*, 1996). Other possible candidate genes excluded on the basis of their localisation include plakoglobin and two members of the Discs-Large gene family, DLG2 and DLG3 (Aberle *et al*, 1995; Mazoyer *et al*, 1995; Smith *et al*, 1996). Plakoglobin is a desmosome-associated protein with homology to the catenin family implicated in cell-cell adhesion modulation (reviewed Garrod, 1995). DLG2 and DLG3 are also thought to have a role in cell-cell adhesion and cell membrane organisation. Mutations in the

Drosophila homologue of DLG2 have been shown to impair cell-cell adhesion and result in proliferation of the imaginal disc epithelium (Woods and Bryant, 1989).

6.12(i) Somatic cell hybrid characterisation of the chromosomal region 17q24-25

DNA from a series of human/rodent somatic cell hybrids containing fragments of chromosome 17 (Black *et al*, 1993) were characterised with microsatellite marker loci mapping around the TOC locus. The aim of this approach was to determine if these hybrids may be useful in localising candidate genes to the region harbouring the disease locus. These hybrids have been well characterised with DNA marker loci mapping in the 17q12-21 region, but only poorly characterised for more telomeric DNA marker loci (Black *et al*, 1993; Prof. E. Solomon, personal communication). Three hybrids were chosen which had maintained fragments of the human complement of chromosome 17 telomeric to 17q12-21: PLT8, TLT8 and TLT10 (gift from Prof. E. Solomon). DNA extracted from these hybrid DNA samples were PCR amplified with microsatellite marker loci localised in the chromosomal region harbouring TOC (as described in section 2.6.1). An ideogram summarising the presence/absence of a locus-specific PCR product in each hybrid is shown in Figure 6.7 (a). Examples of ethidium bromide stained 2 % agarose gels of PCR fragments derived from the hybrid DNA screen are shown in Figure 6.7 (b). The results generated are consistent with the genetic order of the microsatellite marker loci (see section 6.9).

The microsatellite marker loci fall into three groups based on the 17q hybrid DNA content:

- (i) the most centromeric microsatellite marker locus, D17S1826, is absent in all three hybrids tested (gel data not shown)

Figure 6.7. 17q24-25 characterisation of somatic cell hybrid DNA

Analysis of DNA extracted from somatic cell hybrids which contain specific fragments of chromosome 17 (Black et al, 1993). These were screened for presence or absence of microsatellite marker loci which span the TOC locus by PCR amplification (as described in section 2.6.1).

(a) Graphic representation of results from PCR amplification of the hybrid DNA panel

The order of microsatellite marker loci are shown positioned along the single line, including their orientation with respect to the chromosome 17 centromere (CEN) and telomere (TEL). The minimal region harbouring the TOC locus is shown to the left of the marker loci. The somatic cell hybrids are indicated at the top of the figure. Arrowed lines underneath each hybrid indicate the presence of specific marker loci in each hybrid. PCTBA1.8 is a monochromosomal hybrid and contains the whole of chromosome 17 (Kelsell *et al* , 1995a).

(b) Examples of PCR amplification of the hybrid DNA panel.

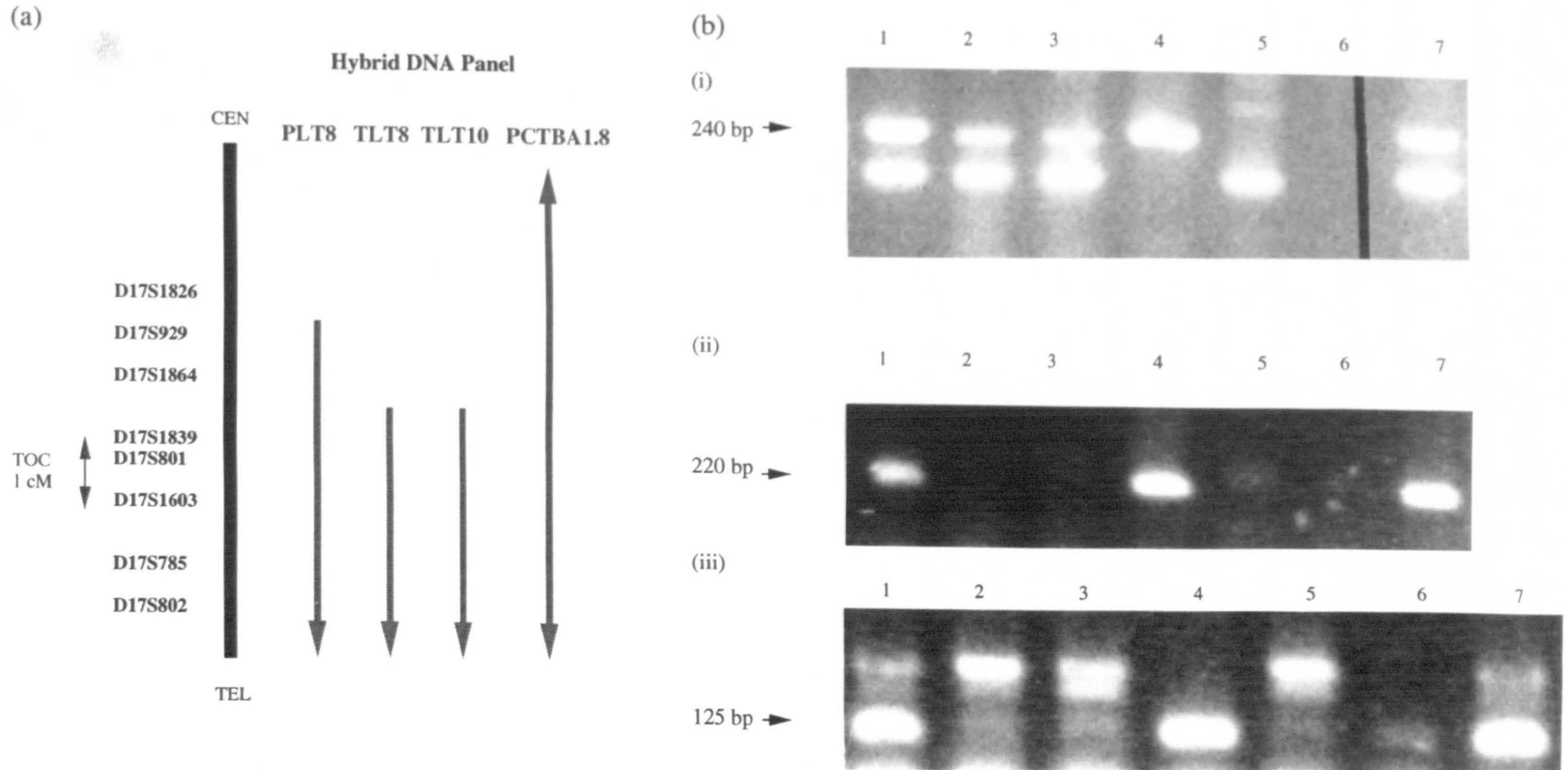
Photographs of 2 % agarose gels showing DNA products from the PCR amplification of DNA extracted from somatic cell hybrids and species-specific DNA templates. The arrow to the left of each gel indicates the human-specific PCR amplification product (sized previously as described in section 2.7.3). The lane order is indicated above each gel and the loading order is the same for each gel. The DNA template added to the PCR reaction loaded in each lane is as follows: Lane 1: PLT8, lane 2: TLT8, lane 3: TLT10, lane 4: human, lane 5: mouse, lane 6: water, lane 7: PCTBA1.8.

(i) D17S1839 specific-PCR amplification products are present in lanes: 1, 2, 3, 4 and 7. The smaller sized DNA fragment is mouse-specific as it is present in lanes: 1, 2, 3, 5, 7. It is clearly distinguishable from the human D17S1839 PCR product.

(ii) D17S929 specific-PCR amplification products are present in lanes: 1, 4 and 7.

(iii) D17S1864 specific-PCR amplification products are present in lanes: 1, 4 and 7. The larger sized DNA fragment is mouse-specific as it is present in lanes: 1, 2, 3, 5, 7. It is clearly distinguishable from the human D17S1864 PCR product.

Figure 6.7. 17q24-25 characterisation of somatic cell hybrid DNA



(ii) the most telomeric group of microsatellite marker loci: D17S1839 (Figure 6.7.b.(i)), D17S801 (Figure 6.8 (b)), D17S1603 (Figure 6.8 (a)), D17S785 (gel data not shown) and D17S802 (gel data not shown) are present in all three hybrids

(iii) microsatellite marker loci mapping between D17S1826 and D17S1839: D17S929 and D17S1864 are only present in hybrid PLT8 (Figure 6.7.b.(ii) and (iii))

This mapping data suggests the hybrid panel has limited use in assigning transcribed sequences to the TOC minimal region. Its main limitation is the inability to localise loci centromeric to D17S1603 or indeed, D17S802 (which is a genetic distance of 4 cM telomeric to D17S801: Dib *et al.*, 1996). Therefore, the characterisation of a physical contig of the region is a priority for the localisation of candidate genes for TOC (described in the next section).

6.12(ii) Towards a YAC (and BAC) contig of the TOC region

Haplotype analysis of focal NEPPK/oesophageal families has defined an approximately 1 cM minimal interval in which the TOC gene is located (section 6.9: Kelsell *et al.*, 1996b). This genomic region is of a small enough size to suggest the construction of a physical map, using YACs situated between the flanking marker loci D17S1839 and D17S1603, is feasible (compared to 6 cM for BRCA2: see section 4.3).

Three YACs were identified on the basis of their STS content from a published integrated physical and genetic map of chromosome 17 (Chumakov *et al.*, 1995: as described in section 2.11.1). Two YAC clones were selected which contain the microsatellite marker locus, D17S801: YACs 786_f_7 and 868_g_7. One YAC clone was selected which contains the microsatellite marker locus, D17S1603: YAC 644_b_10 (Figure 6.8 (a) lane 7). However, two lines of investigation demonstrated that all three of the YACs were chimeric and contained co-ligated human DNA inserts mapping to distinct regions of the genome:

(i) fluorescent *in situ* hybridisation of total YAC DNA to human lymphocyte metaphase spreads revealed hybridisation signals on other chromosomal regions including 17q24-25 for all three YACs (S. Ford, personal communication).

(ii) *Alu* fingerprint DNA profiles of human DNA inserts from the YACs 786_f_7, 868_g_7 and 644_b_10 indicated "shared fragments" with human DNA inserts from YACs which contain loci mapping to other regions of the genome (as described in section 2.11.1). YAC 786_f_7 shares *Alu* DNA fragments with, for example, the non-chimeric YACs 730_a_4 and 786_f_12 which contain the microsatellite marker loci D8S1077 and D12S353, respectively (data not shown). The high degree of chimerism demonstrated by these YACs makes them unsuitable as a starting material for the construction of a physical contig of the TOC region. No YAC clones were identified for the microsatellite marker loci D17S1864 and D17S1839 (as described in section 2.11.2).

Microsatellite marker loci spanning the TOC locus were then used to screen a Bacterial Artificial Clone (BAC) library (as described in section 2.12). Bacterial Artificial Chromosomes (BACs) are vectors based on the low-copy number replicon found in the F-factor system of *E.coli* which can harbour cloned DNA fragments of around 300 kb (Shizuya *et al.*, 1992; Strachan and Read, 1996). Three BAC clones (149.14.1, 88.6.13, and 149.14.3) were identified which contain the telomeric flanking microsatellite marker locus for TOC, D17S1603 (Figure 6.8 (a)). DNA extracted from these BAC clones (as described in section 2.3.5) were screened by PCR amplification for presence or absence of the nearest centromeric microsatellite marker locus to D17S1603, D17S801 (as described in section 2.6.1). The absence of D17S801 PCR amplification products in these BAC clones (Figure 6.8 (b)) indicated that these two microsatellite marker loci were not present in the same BAC clone. Fluorescent *in situ* hybridisation of total BAC DNA to human lymphocyte metaphase spreads revealed hybridisation signals only at 17q24-25 (S.Ford, personal communication). This indicates that these BACS are unlikely to be chimeric and provide a starting point for the construction of a physical map of the TOC region. This work is ongoing.

Figure 6.8. Somatic cell hybrid, YAC and BAC characterisation with D17S1603 and D17S801.

This figure shows the analysis of DNA extracted from somatic cell hybrids which contain specific fragments of chromosome 17 (Black et al, 1993); the YAC 644_b_10 and three BAC clones (149.14.1, 88.6.13 and 149.14.3). These were screened for presence or absence of the microsatellite marker loci, D17S1603 and D17S801 by PCR amplification (as described in section 2.6.1). (An ideogram showing the 17q24-25 DNA content of the somatic cell hybrids is shown in Figure 6.7(a)).

(a) D17S1603

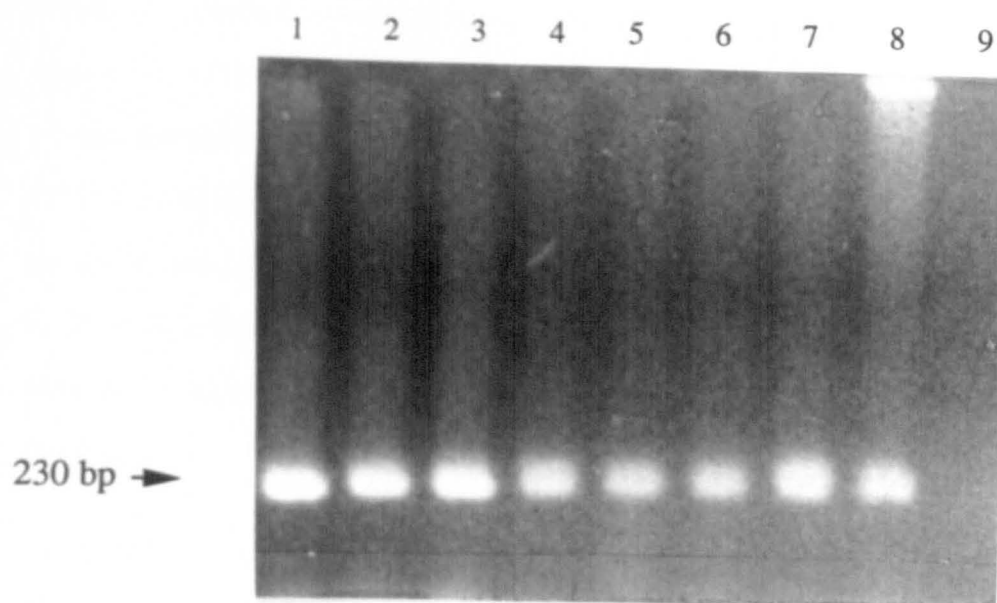
2 % agarose gel showing products from the PCR amplification of DNA extracted from somatic cell hybrids, a YAC, three BAC clones and species-specific DNA templates with D17S1603. The arrow to the left of each gel indicates the D17S1603 PCR amplification product (sized previously as described in section 2.7.3). The lane order is indicated above the gel image. The DNA template added to each PCR amplification reaction loaded in each lane is as follows: Lane 1: BAC 149.14.1, lane 2: BAC 88.6.13, lane 3: BAC 149.14.3, lane 4: hybrid PLT8, lane 5: hybrid TLT8, lane 6: hybrid PLT10, lane 7: YAC 644_b_10, lane 8: human, lane 9: mouse. D17S1603 PCR amplification products are present in lanes: 1, 2, 3, 4, 5, 6, 7, and 8.

(b) D17S801

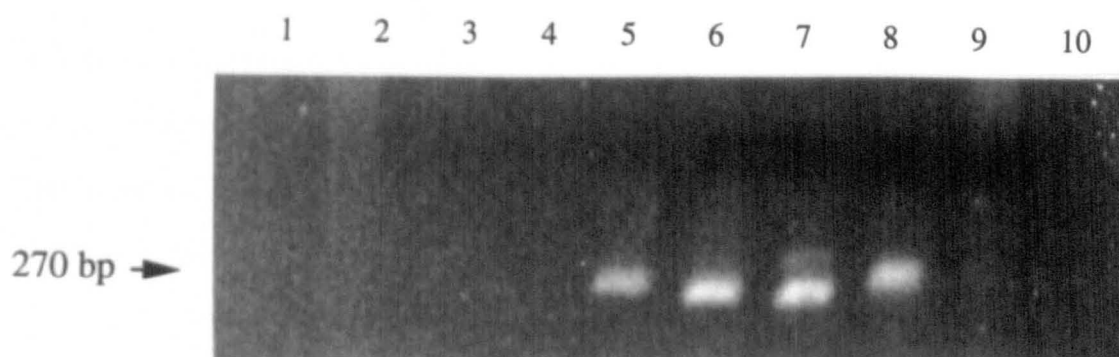
2 % agarose gel showing products from the PCR amplification of DNA extracted from somatic cell hybrids, a YAC, three BAC clones and species-specific DNA templates with D17S801. The figure is as described above for Figure 6.8(a). The DNA template added to the PCR amplification reaction loaded in each lane is as follows: Lane 1: BAC 149.14.1, lane 2: BAC 88.6.13, lane 3: BAC 149.14.3, lane 4: YAC 644_b_10, lane 5: hybrid PLT8, lane 6: hybrid TLT8, lane 7: PLT10, lane 8: human, lane 9: mouse, lane 10: no DNA. D17S801 PCR amplification products are present in lanes: 5, 6, 7 and 8. The differences in product sizes are due to variations in the number of (CA) repeat blocks.

Figure 6.8. Somatic cell hybrid, YAC and BAC characterisation with D17S1603 and D17S801.

(a) D17S1603



(b) D17S801



6.12(iii) Envoplakin, a candidate gene for TOC

A 210 kD protein, envoplakin (EVPL), which demonstrates significant homology to the intermediate filament-binding proteins desmoplakin and plectin, has been localised to 17q25 (C. Ruhrberg, ICRF Keratinocyte laboratory, Lincoln's Inn Fields, London: personal communication). Like desmoplakin and plectin, a series of heptad repeats are found in the central domain of EVLP suggesting it may exist as a coiled-coil dimer. EVLP is expressed in stratified squamous epithelium from the oral mucosa and the oesophagus; but its expression in palmoplantar epidermis has yet to be examined (C. Ruhrberg, personal communication). The EVLP protein shows partial cellular localisation to the desmosomes and may link keratin filaments to the desmosomes which are bound to the cornified envelope (Figure 6.9a). If EVLP specifically binds to KRT16 (or its type II keratin partner), a mutated copy of EVLP could feasibly cause a similar phenotypic disruption to epidermal cells as KRT16 affected individuals with focal NEPPK. Therefore, EVLP could be regarded as a candidate gene for TOC on the basis of its chromosomal location, pattern of tissue expression and its putative function.

Using gene specific oligonucleotide primers for EVLP, the YAC 644_b_10 was identified from a PCR amplification screen of the CEPH YAC library (C. Ruhrberg, personal communication). This YAC also contains the telomeric flanking microsatellite marker locus for TOC, D17S1603 (discussed in previous section). FISH localisation of a cosmid clone containing EVLP to 17q25 (C. Ruhrberg, personal communication) supports EVLP mapping adjacent to D17S1603.

EVLP-specific oligonucleotide primers (as described in section 2.6.7) were used to screen DNA extracted from the panel of chromosome 17 hybrids, the YAC 644-b_10 (extracted as described in section 2.3.4) and the three BACs identified with D17S1603. EVLP-specific PCR products were detected in the same hybrids as for D17S1603 (see ideogram of hybrid content, Figure 6.7 (a)) supporting the localisation of EVLP to the TOC region (Figure 6.9 (b)). However, no EVLP PCR products were amplified when the DNA template was from the three BACs containing D17S1603.

Figure 6.9. Envoplakin (EVLP), a candidate for the TOC gene?

(a) Putative role of EVLP in cornified envelope formation.

Simplified model for a proposed role for EVLP in linking keratin filaments to the desmosomal junctions (adapted from C. Ruhrberg and Dr. F. Watt, personal communication).

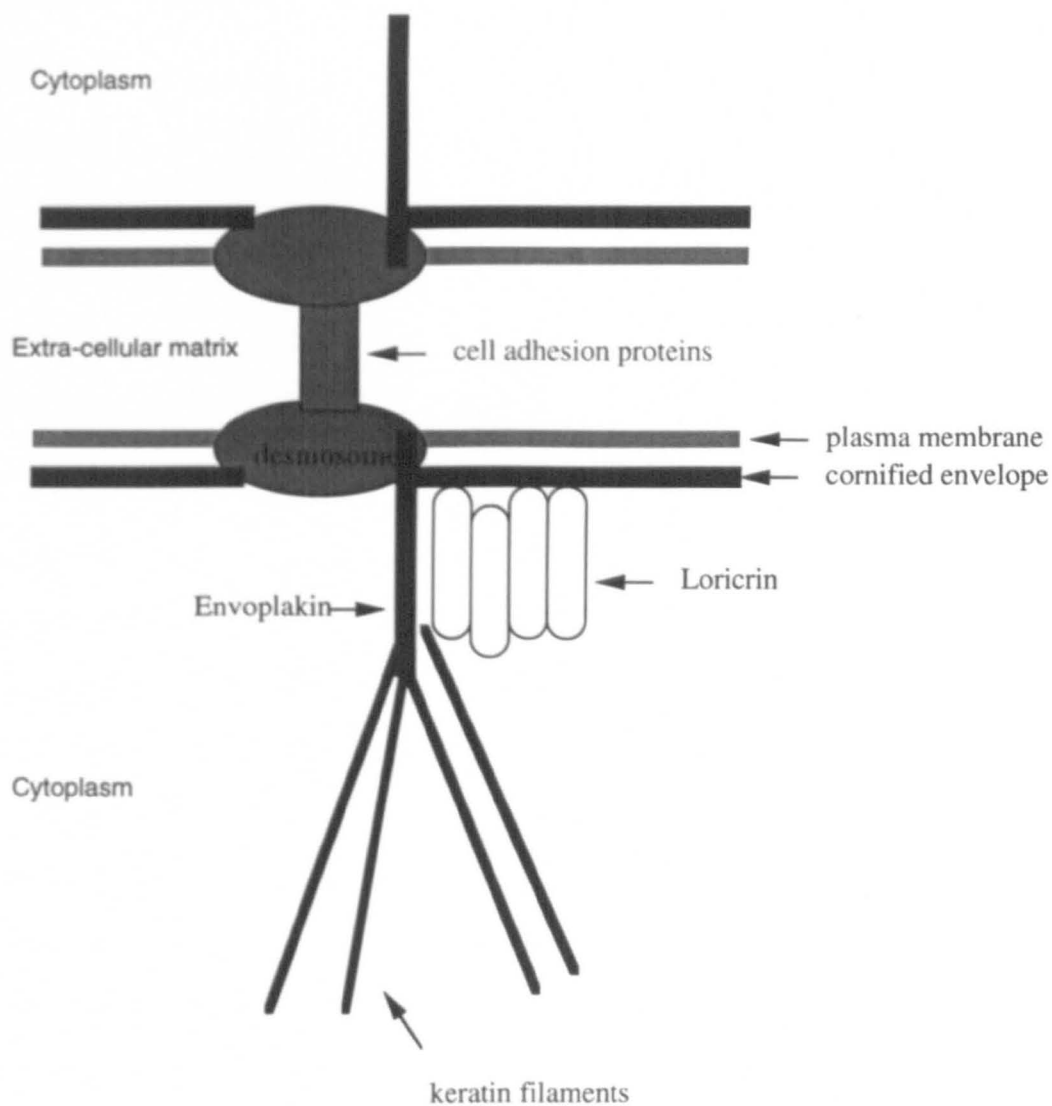
(b) Confirmation of the assignment of EVLP to the TOC region

2 % agarose gel showing products from the PCR amplification of DNA extracted from somatic cell hybrids, a YAC, three BAC clones, human and rodents with human EVLP oligonucleotide primers. The arrow to the left of the gel indicates the EVLP PCR amplification product. The lane order is indicated above the gel image. Lane 1 ϕ x174 DNA digested with *Hae*III molecular weight marker (Boehringer Mannheim).

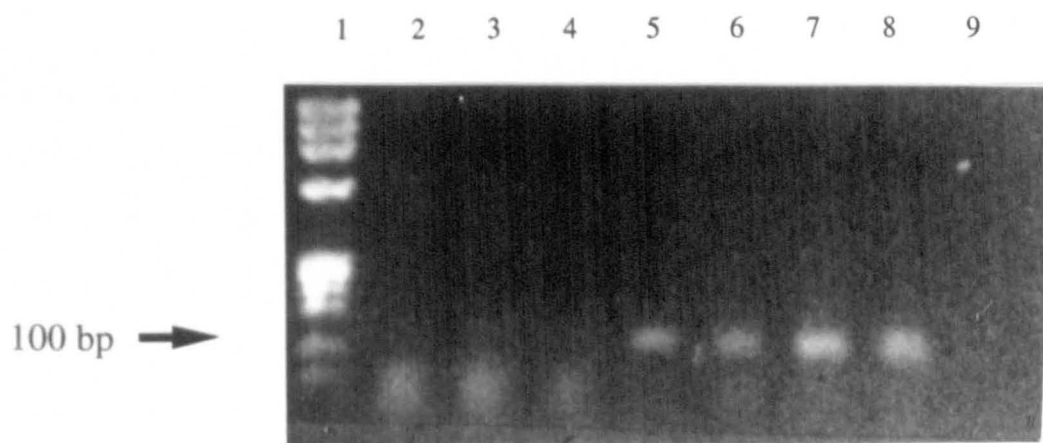
The DNA template added to each PCR amplification reaction loaded in each lane is as follows: Lane 2: BAC 149.14.1, lane 3: BAC 88.6.13, lane 4: BAC 149.14.3, lane 5: hybrid PLT8, lane 6: hybrid TLT8, lane 7: hybrid PLT10, lane 8: YAC 644_b_10, lane 9: mouse. EVLP PCR amplification products are present in lanes: 5, 6, 7, and 8.

Figure 6.9. Envoplakin (EVLP), a candidate for the TOC gene?

(a). Putative role of EVLP in cornified envelope formation



(b) Assignment of EVLP to the TOC region



Further work is in progress to determine the orientation of EVLP to D17S1603. If EVLP maps centromeric to D17S1603, it can be regarded as a strong candidate for the focal NEPPK/oesophageal cancer gene, TOC.

6.13 Summary of TOC analysis

The TOC locus has been mapped to a 1 cM region flanked by D17S1839 and D17S1603 (Kelsell *et al*, 1996b). From the haplotype comparison study described in section 6.10, two disease haplotypes were identified in the three families analysed. The two largest families, UK and American, have different haplotypes segregating with the disease suggesting two distinct mutations in the same gene. This may explain the difference in oesophageal cancer risk between the two families, with the UK mutation conferring an elevated risk of cancer in TOC gene carriers.

The high incidence of oesophageal cancer observed in the TOC-PPK affecteds compared to the KRT16-PPK affecteds is interesting as they present with similar patterns of keratoderma. This may be due to differences in the function of the proteins encoded by the two distinct genetic susceptibility loci in cytoskeletal organisation. It is possible that the TOC gene may anchor the KRT16 filaments to the cornified envelope and therefore, if mutated, may increase the permeability of the cell to environmental carcinogens. A gene, envoplakin (EVLP), which may possibly link keratins to the cornified envelope and maps in the region harbouring the TOC gene, is currently under investigation as a strong candidate for the genetic basis of this disease. This is discussed further in Chapter 7.

6.14 Punctate NEPPK

Punctate NEPPK presents with punctate keratoses over the surface of the palm and sole (Figure 6.1 (c)). A four generation family, 2039, with punctate NEPPK has been identified (Figure 6.10; Stevens *et al*, 1996b). The disease appears to be inherited as an autosomal dominant trait presenting with phenotypically identical multiple keratoses in affected family members. All keratoses developed from the age of puberty, though

the age of onset varied from 13 to 40 years of age. In addition, the number and size of keratoses varied between family members, some only noticeable after immersion in water (Dr. H. Stevens, personal communication). A possible association between PPK and cancer was noted in the family, with a number of tumours developing before the age of 50 (Stevens *et al*, 1996b). DNA extracted from the lymphocytes of affected and unaffected family members was genotyped with microsatellite marker loci spanning the two keratin clusters and other candidate gene regions (as described in section 2.7). Linkage analysis was performed as described in section 2.13.2.

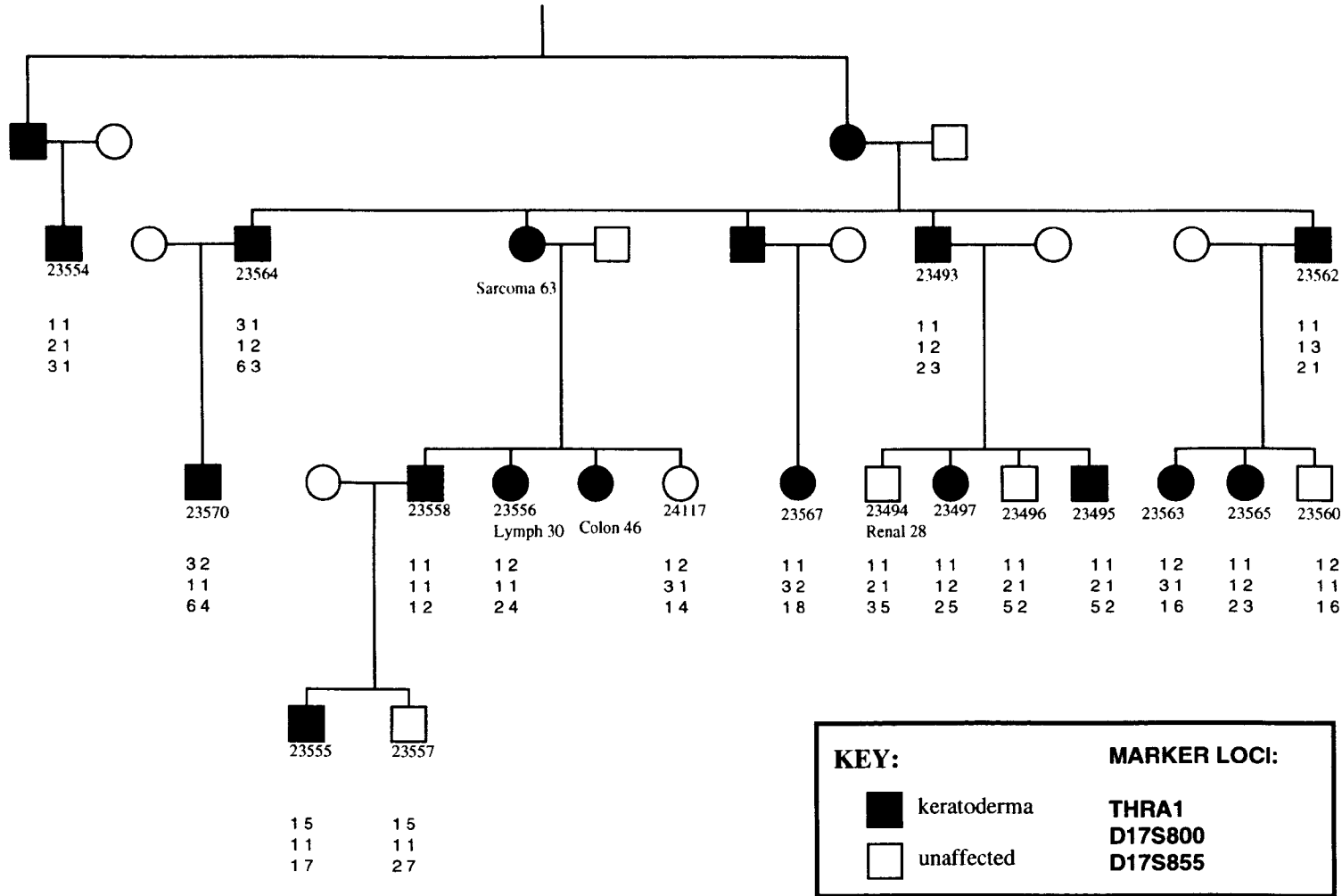
Close linkage to the disease was excluded with microsatellite marker loci mapping near to the epithelial keratin gene clusters on 12q11-13 and 17q12-21. A LOD score of -2.3 at a recombination fraction of 0.1 was obtained with the marker D17S855 and -3.04 at a recombination fraction of 0.00 with the marker D12S368 (Table 6.6). The lack of a 17q12-21(or 12q11-13; data not shown) haplotype in common between affected family members further supports the exclusion of a type I (or type II) keratin as the genetic basis of this disease (Figure 6.10). Evidence against linkage of the disease to other regions of the genome implicated in hereditary skin diseases was also demonstrated including: D12S330 (Darier's disease: Parfitt *et al*, 1994), D17S801 (TOC: this study) and D3S1292 (Hailey-Hailey disease: Ikeda *et al*, 1994) (Table 6.6). The genetic basis of punctate NEPPK and possibly cancer susceptibility requires further studies.

Table 6.6. Lod scores obtained for pedigree 2039: evidence against linkage with punctate PPK and candidate chromosomal regions								
	0.0	0.01	0.05	0.1	0.2	0.3	0.4	
D17S855	-5.05	-4.02	-2.91	-2.30	-1.75	-1.16	-0.47	Type I keratins
D12S368	-3.04	-1.57	-0.34	0.06	0.24	0.17	0.05	Type II keratins
D12S330	-5.48	-4.05	-2.39	-1.48	-0.62	-0.23	-0.05	Dariers disease
D17S801	-3.21	-3.16	-2.33	-1.44	-0.58	-0.20	-0.04	17q24-25 (TOC)
D3S1292	-3.72	-3.09	-1.78	-1.13	-0.49	-0.19	-0.04	Hailey-Hailey disease

Figure 6.10. Genetic exclusion of the type I keratins with punctate palmoplantar keratoderma segregating in pedigree 2039.

The sex of each individual is as described in Figure 1.1, the identification number is shown underneath each individual. The type of malignancy with age of diagnosis is given beneath the affected individual. PPK affected status is as indicated from the key in the lower right hand corner of the figure. The order of microsatellite marker loci is also shown to the bottom right of the figure (see also Figure 3.10). Detection of alleles for each microsatellite marker locus is as described in section 2.7.1. Haplotypes are displayed under each individual typed. No specific 17q12-21 haplotype is segregating with punctate PPK in this pedigree.

Figure 6.10 Genetic exclusion of the type I keratins with Punctate palmoplantar keratoderma segregating in pedigree 2039



6.15 Summary of Chapter 6

In this chapter, the genetic analysis of families with three forms of non-epidermolytic palmoplantar keratoderma (NEPPK): focal, diffuse and punctate; was performed. In some of the pedigrees, a striking association of focal NEPPK with oesophageal cancer susceptibility was observed. This study demonstrated genetic heterogeneity between the three forms of NEPPK (Kelsell *et al*, 1995b).

Genetic heterogeneity was also established between families presenting with clinically similar forms of focal NEPPK. Mutations in the type I keratin, KRT16, were identified as the genetic basis of focal NEPPK in a pedigree without associated susceptibility to oesophageal cancer (Shamsher *et al*, 1995). Haplotyping of three families with focal NEPPK associated with elevated susceptibility to oesophageal cancer has excluded mutations in KRT16 as the genetic basis of disease in these families. Disease in these families is genetically linked to a 1 cM region on chromosome 17q24-25; the TOC locus (Kelsell *et al*, 1996b).

CHAPTER 7. DISCUSSION

7.1. Introduction

This study focussed on two areas in the field of cancer susceptibility. The initial focus was the genetic analysis of a recently mapped breast cancer susceptibility locus, BRCA1, in a number of breast and breast-ovarian cancer families (Chapters 3, 4 and 5). During the course of this investigation, there have been dramatic developments in breast cancer genetics culminating in the identification of two major breast cancer susceptibility genes, BRCA1 and BRCA2 (Miki *et al*, 1994; Wooster *et al*, 1995; Tavtigian *et al*, 1996). The other area of study has been the genetic analysis of a group of autosomal dominant skin diseases, termed the non-epidermolytic palmoplantar keratodermas (NEPPK: Chapter 6). Although the skin disorders themselves are only mildly debilitating, there is an association of the skin disease with the development of malignancy in a number of pedigrees. Indeed in one family, the risk of dying from oesophageal cancer for an individual affected with NEPPK has been calculated to be 92 % by age 70 years (Ellis *et al*, 1994).

7.2. Breast-ovarian cancer

7.2.1. Linkage study

The large breast-ovarian cancer family, BOV3, provided the majority of evidence for linkage of cancer susceptibility to the locus BRCA1 (Chapter 3). Initial linkage studies of the other 15 families gave little information towards confirming BRCA1 association and narrowing the genetic region harbouring this locus. This was because of problems often encountered in mapping common disease genes ie. heterogeneity, a partially penetrant gene, phenocopies, small pedigrees and limited DNA material, particularly from affected individuals. For example, an individual with ovarian cancer, which did not share the putative disease haplotype segregating with the disease in the rest of the BOV3 family, suggested disease in this family may not be linked to BRCA1. However, analysis of further genetic marker loci and additional affected individuals in this family clarified the situation and demonstrated this individual was likely to be a case of “sporadic” disease (Chapter 3: Section 3.5 (a)). All linkage data generated from

families genotyped worldwide was analysed by the Breast Cancer Linkage Consortium. This collaborative study demonstrated the power of analysing a large data set for such a complex disease (Easton *et al*, 1993b). Analysis of 57 breast-ovarian cancer families suggested that disease in the majority was likely to be due to an inherited mutation in BRCA1. Subsequent to this analysis, individual families could be reassessed as to their likelihood of disease being linked to BRCA1 and contradictory locations for BRCA1 (from individual recombination events) could be interpreted with greater clarity.

Analysis of the BOV3 family was the main focus for this part of the study resulting in a number of findings, including the identification of a recombination event which significantly reduced the BRCA1 minimal region to an interval of 1-1.5 Mb (Chapter 3; Section 3.5; Kelsell *et al*, 1993). After the cloning of BRCA1 (Miki *et al*, 1994), a 185delAG BRCA1 mutation in this family was identified (Shattuck-Eidens *et al*, 1995; Xu *et al*, in press). The same mutation was also identified in two other ICRF families, ICRF178 and ICRF546 (Neuhausen *et al*, 1996a; Xu *et al*, in press). Haplotype analysis revealed that gene carriers from these two families had inherited the same ancestral mutation as the BOV3 family (Chapter 5: Section 5.3(iv)). These three families were all ascertained in the north of England. Disease haplotypes in gene carriers from these families were extended by genotyping an additional 12 microsatellite marker loci spanning a genetic distance of, approximately, 47 cM on the long arm of chromosome 17. As approximately 23 cM of the putative ancestral disease haplotype was in common, this would suggest the three families are closely related, separated by only a few generations, even though family members have no knowledge of any genetic links between the three families. Further studies are in progress to determine the relationship between these three families. Founder mutations have been identified which frequently occur in two "isolated" populations: in Iceland (BRCA2:Thorlacious *et al*, 1996) and in the Ashkenazi Jewish population (BRCA1 and BRCA2: Struwing *et al*, 1995a; Neuhausen *et al*, 1996b). Interestingly, these latter two founder mutations (one in BRCA1, the other in BRCA2) are likely to

account for a quarter of all early-onset breast cancer cases in the Ashkenazi population (Neuhausen *et al*, 1996b).

Compelling evidence from the analysis of allele status in DNA extracted from breast tumours of the BOV3 family implicated BRCA1 as a tumour suppressor gene (Chapter 5: Section 5.2). Breast tumour DNA samples from nine gene carriers in this family were hemizygous for the mutated BRCA1 allele. Further support came from the BCLC study in which 86 % of breast and ovarian tumours from putative BRCA1 gene carriers also demonstrated loss of the wildtype BRCA1 allele (Cornelis *et al*, 1995). On the strength of these observations, allele loss data was incorporated into the linkage analysis of disease association with BRCA1 in 63 breast and breast-ovarian cancer families (Chapter 5: Section 5.3). This was found to be extremely helpful in analysing pedigrees of small size and/or limited sample availability to improve LOD scores and assess whether disease in these families was likely to be due to a germline BRCA1 mutation.

The same set of families genotyped with microsatellite marker loci flanking BRCA1 (Chapter 5: Section 5.3) were also assessed for linkage of the disease to BRCA2 (E. Mavraki, personal communication). This was to estimate, by multipoint lod score analysis, the proportion of families in which disease was likely to be due to BRCA1 or BRCA2 and to assess if there is evidence for a presently unidentified third locus for breast-ovarian cancer susceptibility, BRCA3 (Prof. T. Bishop and G. Crockford, personal communication). From this data set, statistical analysis suggests that disease in the majority of families is likely to be due to either a mutation in BRCA1 or BRCA2 in approximately equal proportions. However, this analysis does not preclude the possibility of the existence of BRCA3 or further breast cancer susceptibility loci. This is due to the fact that some families studied added only a limited amount of information (that is, LOD scores close to zero for each locus) to the analysis due to their small size (in terms of structure and sample availability). Therefore, it is possible that BRCA3 (or BRCA4 etc) could account for disease susceptibility in a proportion of these pedigrees. It is also likely that cases of sporadic disease in these families, particularly those of

limited sample availability, may be affecting the interpretation of the linkage and haplotype analysis in these pedigrees. For example, a negative lod score at BRCA1 (or BRCA2) with the disease in a family does not exclude the possibility that disease, in a proportion of cases in that family, is in fact due to a BRCA1 (or BRCA2) germline mutation (for example, Friedman *et al*, 1995; Phelan *et al*, 1996). A more accurate assessment of the proportion of families due to BRCA1 and BRCA2 would require the combination of linkage analysis with DNA mutation analyses of both susceptibility genes in lymphocyte DNA extracted from a number of affected family members. Mutation analyses are currently being performed in the 63 ICRF families which had been previously assessed for linkage with the disease to BRCA1 or BRCA2 (BRCA1: Prof. E.Solomon, personal communication; BRCA2: Elena Mavraki, personal communication-see next section).

Putative gene carriers from six ICRF breast cancer families (out of the 63 families analysed) showing supportive evidence of linkage of the disease to BRCA2 (Prof. T. Bishop, personal communication) have been screened for mutations in BRCA2 (E. Mavraki, personal communication). After screening the whole of the coding region of BRCA2, a germline BRCA2 mutation has been identified in one of the six families. A recent report identified germline BRCA2 mutations in 8 of 49 site-specific breast cancer families (Phelan *et al*, 1996). These two studies would suggest that BRCA2 does not account for such a high proportion of familial breast cancer cases as previously estimated from linkage studies. The relatively low proportion of families with identified BRCA2 mutations is in contrast to the much higher proportion of breast and breast-ovarian cancer families with identified BRCA1 mutations. For example, BRCA1 mutations were identified in 26 of 37 families analysed (Friedman *et al*, 1995) and 22 of 32 from another study (Gayther *et al*, 1995).

7.2.2. Role of BRCA1 and BRCA2 in sporadic breast cancer

The roles of BRCA1 and BRCA2 as classical tumour suppressor genes in sporadic breast cancer is at present unclear. The relatively high number of sporadic breast

tumours displaying allele loss at BRCA1 (47 %: Chapter 5: Section 5.4) and BRCA2 (45 %: Chapter 5: Section 5.5) suggests a role for these genes in sporadic disease. However, no somatic BRCA1 or BRCA2 mutations have yet been identified in sporadic breast carcinomas (for example, Futreal *et al*, 1994; Lancaster *et al*, 1996). (No somatic BRCA1 or BRCA2 mutation analyses have been performed in the sporadic breast tumours assessed for allele loss in this study). A clue may lie in the observations that breast tumours from BRCA1 gene carriers often have the histology of grade 3 infiltrating ductal carcinomas (for example, Eisinger *et al*, 1996). This tentatively suggests a role for BRCA1 in sporadic disease with this histology. However, allele loss at BRCA1 in sporadic breast carcinomas demonstrated no correlation with histological grade (statistical analysis performed by Dr. D. Barnes and colleagues: data not shown).

From the analysis of somatic events occurring in grade 3 infiltrating ductal carcinomas from BRCA1 gene carriers of the BOV3 family, allele loss at both BRCA1 and BRCA2 was unequivocal in five of the seven tumours analysed (Chapter 5: Section 5.6). Interestingly, re-examination of the allele loss patterns at these two loci in the sporadic breast tumours revealed a similar pattern of combined loss (or retention) of BRCA1 and BRCA2 in those classified as grade 3 infiltrating ductal carcinomas (Kelsell *et al*, 1996a). It is possible that the apparently "sporadic" grade 3 infiltrating ductal carcinomas may be carrying a germline mutation in the retained allele of BRCA1. Alternatively, this may suggest a key role for BRCA1 (and BRCA2) in the development of sporadic tumours of this histology and possibly a compensatory effect between different pathways involving BRCA1 and BRCA2 in breast epithelial cells.

However, as somatic BRCA1 or BRCA2 mutations are extremely rare in breast tumours with (or without) allele loss, distinct mechanisms of gene inactivation may be occurring in sporadic tumours compared to the familial breast tumours. It is possible that loss of one copy of both genes is sufficient for the development of these sporadic breast tumours: a dosage effect. Alternatively, the retained wildtype copies of BRCA1

and BRCA2 may be transcribed at low levels or even silenced. This could be due to methylation of the retained copy of both BRCA1 and BRCA2. Alternatively, somatic mutations may be present in a gene(s) which affects the transcriptional or translational regulation of BRCA1 and BRCA2. It has been demonstrated that BRCA1 mRNA levels are reduced in sporadic breast tumours (Thompson *et al*, 1995). BRCA1 expression also appears to be hormonally regulated by oestrogens and is at its highest level during pregnancy (Marquis *et al*, 1995; Gudas *et al*, 1995). The inhibition of breast epithelial cell growth by hormonally-regulated BRCA1 supports the suggestion, from epidemiological studies, of a protective effect of women having an early pregnancy from breast cancer. A hypothetical model of the role of BRCA1 (and BRCA2) in sporadic versus familial grade 3 infiltrating ductal carcinomas is shown in Figure 7.1.

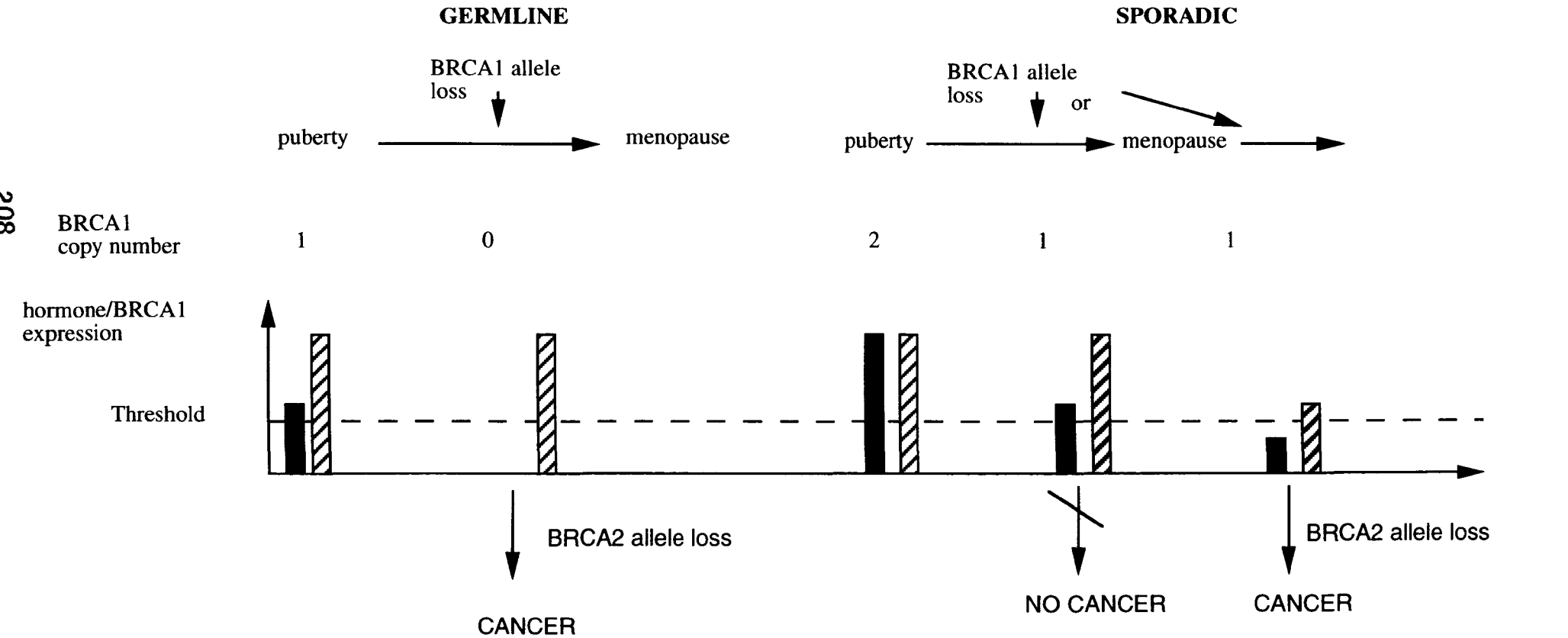
Focussing on the sporadic grade 3 breast tumours, recent analysis (statistical analysis performed by Dr. D. Barnes and colleagues: data not shown) suggests patients with breast tumours displaying combined loss (or retention) of BRCA1 and BRCA2 (chromosomes 17q and 13q, respectively) have a greatly improved prognosis to those with tumours displaying loss of BRCA1 (17q) and not BRCA2 (13q). A possible explanation of this result is that tumours displaying loss at 17q and not 13q may harbour a p53 mutation in the retained allele (as whole chromosome 17 deletions are common: data not shown) and have developed by a distinct tumorigenic pathway. This would result in cells with no functional copy of p53. Indeed, a number of studies have associated somatic p53 mutations in breast tumours with poor prognosis in patients (for example, Borresen *et al*, 1995; Kovach *et al*, 1996). Further studies (by colleagues) of “sporadic” grade 3 infiltrating breast tumours are in progress to explore these preliminary findings. These include:

- 1) BRCA1, BRCA2 and p53 mutation analyses in germline and tumour DNA
- 2) allele loss status at BRCA1 and BRCA2 in a larger group of breast tumours of this grade

Figure 7.1. Hypothetical model for a hormonally regulated dosage effect of BRCA1 (and BRCA2) in sporadic breast cancer

This model is based on the observation of a combined role for BRCA1 and BRCA2 in grade 3 infiltrating ductal carcinomas. It proposes that BRCA1 and BRCA2 either interact directly or along the same pathway in the control of breast epithelial cell proliferation. It also includes the possibility that BRCA1 (and possibly BRCA2) are under hormonal (estrogen) regulation. The levels of estrogen in the model do not account for fluctuations in levels due to menstruation or pregnancy. BRCA1 expression levels are shown as a black bar, estrogen levels as a hatched bar.

208



3) comparing prognosis from various post-cancer therapies within this tumour population

7.2.3. Function of BRCA1

Allele loss studies in tumours from disease gene carriers have implicated BRCA1 as a tumour suppressor gene (for example; Kelsell *et al*, 1993a (Chapter 5); Cornelis *et al*, 1995). A function of normal BRCA1 may be the inhibition of proliferation in mammary epithelial cells (Thompson *et al*, 1995). In support of this, transfection of the wildtype BRCA1 gene inhibited tumour growth and improved the prognosis of MCF-7 derived breast tumours in nude mice (Holt *et al*, 1996). The same study found that near full-length truncated BRCA1 protein inhibited ovarian cancer cell growth but not that of breast cancer cells. This is in agreement with the genotype-phenotype correlations observed previously (Gayther *et al*, 1995). This suggests that the first two-thirds of the BRCA1 protein are functionally important in suppression of ovarian cell proliferation.

Mice deficient in two functional copies of BRCA1 die as early embryos, displaying defects of the neural tube (Gowen *et al*, 1996). This suggests murine BRCA1 protein may be important as a developmental switch for the differentiation of neuroepithelial cells, as well as, later in development, having a role in mammary cell differentiation (Marquis *et al*, 1995). However, in humans, BRCA1 appears to be much more tissue-specific, possibly being activated at puberty to control the differentiation of breast epithelial cells. This is based on the observation of two women who survived to adulthood without a normal copy of BRCA1; in contrast with the BRCA1-deficient mice which died before birth (Boyd *et al*, 1995; Dr. M. Boyd and Dr. D. Black, personal communication: discussed in Chapter 5: Section 5.3 (ii)). These two women, deficient in two functional copies of BRCA1, appeared phenotypically normal and developed breast cancer at comparable ages to female BRCA1 heterozygotes.

The identification of a RING finger (zinc-binding) domain near the amino terminal end of the protein suggests BRCA1 may play a role as a transcription regulator (Miki *et al*, 1994). Recently, it has been proposed that BRCA1 and possibly, BRCA2, may be members of the granin family, suggesting they may encode hormonally-regulated secretory proteins involved in cell membrane signal transduction pathways (Jensen *et al*, 1996). The size of these proteins suggests that BRCA1 and BRCA2 may be multi-functional and could be involved in a number of regulatory pathways in the nucleus, cytoplasm and extra-cellular matrix of breast (and ovarian) epithelial cells.

7.3 Non-epidermolytic palmoplantar keratoderma (NEPPK) and cancer susceptibility

7.3.1. Genetic heterogeneity

The main aim of this investigation was to determine the genetic basis of NEPPK in a number of pedigrees. This study has demonstrated genetic heterogeneity between variant forms of this disease; that is, mutations in different genes can cause similar phenotypes. For example, families with focal NEPPK but no associated increase in cancer risk have a heritable defect in the type I epidermal cytoskeletal intermediate filament protein, keratin 16 (KRT16: Chapter 6: Section 6.6; Shamsher *et al*, 1995). It is also possible that a type II keratin may underlie the genetic basis of diffuse NEPPK (Chapter 6: Section 6.5; Kelsell *et al*, 1995b). Further studies are in progress to identify the specific gene defect for this disease. These findings were not wholly unexpected as keratin mutations have been demonstrated as the genetic basis for a number of disorders displaying epidermal fragility (reviewed Lane, 1994; McLean and Lane, 1995). However, three families with a similar phenotypic form of focal NEPPK to that associated with KRT16 mutations which segregates with squamous cell carcinoma of the oesophagus are not due to a inherited mutation in KRT16, or indeed any of the known epidermal keratin genes. Disease in these three families is linked to the chromosomal region 17q24-25, designated the TOC locus (Chapter 6: Section 6.7 and subsequent sections).

A number of other skin diseases have also been implicated by linkage analyses to be due to genes other than keratins (Darier's disease: Parfitt *et al*, 1994; Hailey-Hailey disease: Ikeda *et al*, 1994; Clouston hidrotic ectodermal dysplasia: Kibar *et al*, 1996). A striated form of PPK is linked to microsatellite marker loci mapping adjacent to the desmosomal cadherin gene cluster on chromosome 18q12 in two families (Hennies *et al*, 1995b; D. P. Kelsell and S. P. Bryant: data not shown). Recently, a mutation in loricrin, a component of the cornified cell envelope, has been demonstrated to be the genetic basis of Vohwinkel's syndrome, a mutilating form of PPK (Maestrini *et al*, 1996). Curiously, this mutation predominantly causes symptoms in the palmoplantar epidermis even though loricrin is expressed in all epidermal cells. A model showing how some of these proteins may interact in forming the cornified cell envelope is shown in Chapter 6: Figure 6.9 (a).

7.3.2. Focal NEPPK and Cancer susceptibility

The genetic basis of focal NEPPK and oesophageal cancer susceptibility has been localised to an approximate 1 cM interval on chromosome 17q24-25; the TOC locus (Chapter 6: Section 6.7 and subsequent sections; Kelsell *et al*, 1996b). Genetic analysis of the three TOC-linked pedigrees suggests that both the skin disease and the cancer susceptibility are likely to be caused by mutations in the same gene from two lines of evidence:

- 1) the two phenotypes segregate together in all three pedigrees, and
- 2) the same ancestral TOC mutation appears to be segregating in two of the pedigrees.

This is in contrast to the epidermolytic PPK pedigree associated with breast-ovarian cancer (Blanchet-Bardon *et al*, 1987; Torchard *et al*, 1994; Friedman *et al*, 1995) in which mutations in two tightly -linked gene loci (KRT9: EPPK and BRCA1: breast-ovarian cancer) are segregating with the two disorders. Although the KRT 16 and TOC form of focal NEPPK cause almost identical hyperkeratotic lesions in the mouth and oesophagus (Dr. H. Stevens and Prof. I. Leigh, personal communication), only TOC gene carriers have a high risk of developing oesophageal cancer. No keratin

mutations to date have been associated with cancer susceptibility. This is probably because they are structural intra-cellular proteins which are indirectly attached to the nuclear and cell membrane. A candidate gene for TOC, envoplakin (EVLP), is a desmosomal-associated protein that putatively links keratin filaments to the cornified cell envelope (C. Ruhrberg and Dr. F. Watt, personal communication). As it may attach to the keratin filaments, a heterozygous mutation in EVLP could confer a similar level of epidermal fragility as a keratin mutation, for example, a mutation in KRT16. EVLP may associate with transmembrane proteins of the desmosome, for example, plakoglobin and cadherins, which appear to be involved in the intra-cellular cascade of response to extra-cellular positional information. Though EVLP is unlikely to have a direct role in the cell cycle or proliferation, a mutation in this gene could plausibly increase the risk of oesophageal cancer in a TOC-gene carrier. For example, when an oesophageal cell carrying a defective copy of envoplakin is physically stressed, this could affect the integrity of the desmosomes leading to a:

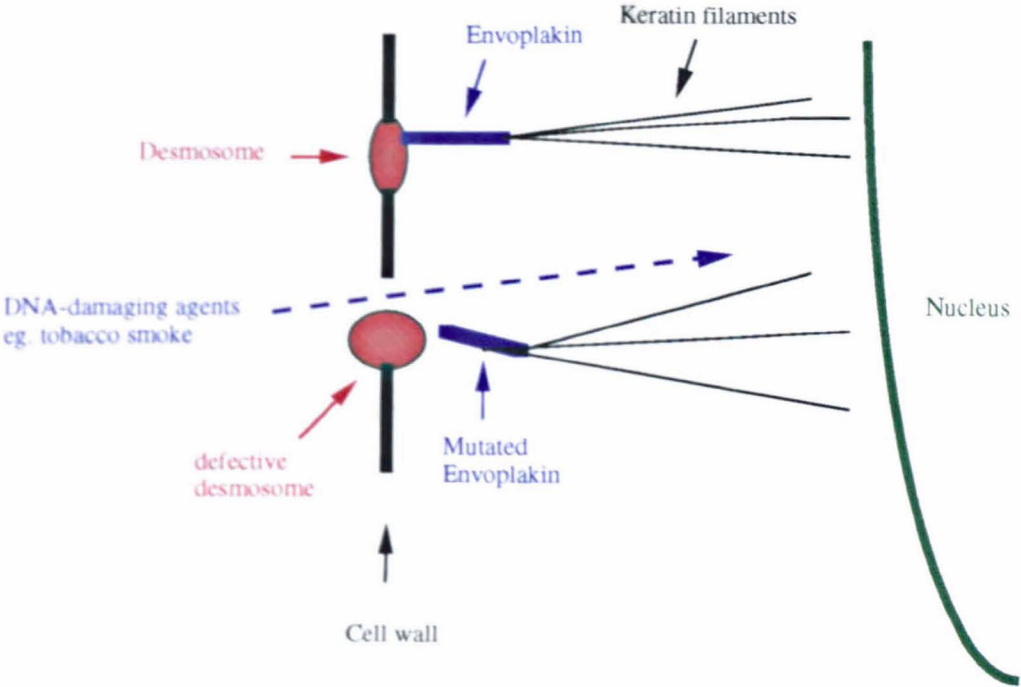
- 1) perturbation of the oral and oesophageal cornified envelope lining causing an ineffective barrier to carcinogen damage, for example, that caused by tobacco smoke, and/or
 - 2) an alteration in desmosomal cell binding affinity may lead to cell isolation as desmosomes are intercellular couplings and mediate cell adhesion. An accumulation of somatic mutations in other genes in such a cell may lead to invasion or metastasis.
- These hypothetical models are shown in Figure 7.2.

Envoplakin mutation analyses of gene carriers from these three families are currently in progress (Dr. J. Risk, C. Ruhrberg and Dr. A. Reis; personal communications). If mutations are identified, Envoplakin can be added to the growing list of cytoskeletal/cell adhesion genes implicated in cancer predisposition or progression. These include the Neurofibromatosis 2 gene (NF2; Trofatter *et al*, 1993), Deleted in Colorectal Cancer gene (DCC: Fearon *et al*, 1990), and the Adenomatous Polyposis Coli gene (APC: Groden *et al*, 1991).

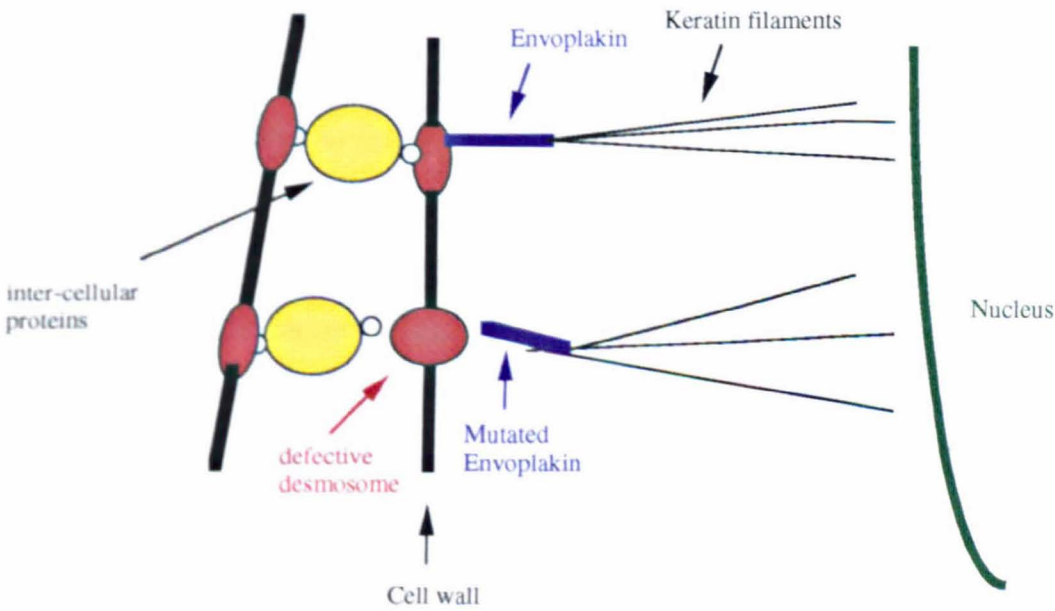
Figure 7.2. Two models implicating EVLP in Oesophageal Cancer susceptibility

Each model depicts two possible scenarios when a cell carrying a mutation in EVLP is stressed.

(a) Increases permeability of cell wall to environmental carcinogen damage



(b) Defective cell adhesion



7.3.3. Punctate PPK and Cancer

Genetic analysis of a family with punctate PPK associated with the development of a range of malignancies including breast cancer was performed (Chapter 6: Section 6.14). A number of candidate gene regions have been excluded as the genetic basis of the skin disease (and cancer susceptibility). An autosomal genome search is currently in progress in this family and others with punctate PPK (Dr. A. Reis, personal communication). Recently, the gene for Cowden disease (or multiple hamartoma syndrome) has been localised to chromosome 10q22-23 (Nelen *et al*, 1996). This is an autosomal dominant disease in which 30 % of affected female individuals develop breast cancer. As one of the phenotypes of Cowden disease is palmoplantar keratoderma, the skin disease and cancer susceptibility in Cowden disease and punctate PPK may have a similar genetic basis.

7.4. Concluding remarks

This study has encompassed a number of the complexities of autosomal dominant gene mapping. For BRCA1, there are the problems associated with mapping a commonly occurring disease such as heterogeneity (the complications caused by BRCA2 and, possibly, an unidentified locus, BRCA3), phenocopies (sporadic cases of breast cancer arising in BRCA1-linked pedigrees) and variable penetrance. Contrastingly, for NEPPK, the disease is much rarer in the population so the problem of phenocopies is minimal. In addition, the disease gene is fully penetrant so gene carriers are easily identifiable. In the cancer associated-forms of NEPPK, an objective was to determine whether the cancer susceptibility was due to a distinct tightly linked locus or if the NEPPK disease gene itself also predisposes individuals to malignancy. For focal NEPPK and associated oesophageal cancer susceptibility (TOC locus), genetic studies support the latter scenario.

Additionally, it is evident that other gene mutations (of low penetrance) and environmental effects may modify the risk and type of cancer in disease gene carriers carrying the same BRCA1 or BRCA2 or TOC mutation. This is apparent from the

study of Icelanders all of whom have inherited the same founder BRCA2 mutation though its penetrance appears extremely variable (Thorlacious *et al*, 1996). By studying these founder populations, it may be possible to identify genes and/or environmental factors that modify cancer susceptibility in BRCA1 or BRCA2 gene carriers.

Finally, from the genetic analysis of rare and often only mildly debilitating inheritable skin diseases, for example, palmoplantar keratoderma, specific proteins that form the complex intra- and inter- cellular interactions of cells can be identified. These proteins may also have a role in sporadic cancers. This will aid in the understanding of cell adhesion defects that underlie a cell's ability to isolate itself and then accrue the genetic alterations required for it to become cancerous.

REFERENCES

- Aaltomaa, S., Lipponen, P. and Eskelinen, M. (1992). Demographic prognostic factors in breast-cancer. *Acta Oncol* 31: 635-640.
- Aberle, H., Bierkamp, C., Torchard, D., Serova, O., Wagner, T., Natt, E., Wirshing, J., Heidkamper, C., Montagna, M. and Lynch, H. T. (1995). The human plakoglobin gene localizes on chromosome 17q21 and is subjected to loss of heterozygosity in breast and ovarian cancers. *Proc Natl Acad Sci U S A* 92: 6384-6388.
- Adams, M. D., Soares, M. B., Kerlavage, A. R., Fields, C. and Venter, J. C. (1993). Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nature Genet* 4: 373-386.
- Adams, M. D., Kerlavage, A. R., Fleischmann, R. D., Fuldner, R. A., Bult, C. J., Lee, N. H., Kirkness, E. F., Weinstock, K. G., Gocayne, J. D., *et al* (1995). Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature*. 377: 3-174.
- Albanes, D. (1987). Caloric intake, body weight, and cancer: a review. *Nutr Cancer* 9: 199-217.
- Albertsen, H. M., Smith, S. A., Mazoyer, S., Fujimoto, E., Stevens, J., Williams, B., Rodriguez, P., Cropp, C. S., Slijepcevic, P. and Carlson, M. (1994). A physical map and candidate genes in the BRCA1 region on chromosome 17q12-21. *Nature Genet* 7: 472-479.
- Anand, R., Villasante, A. and Tyler, S. C. (1989). Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis. *Nucleic Acids Res* 17: 3425-33.
- Arason, A., Barkardottir, R. B. and Egilsson, V. (1993). Linkage analysis of chromosome-17q markers and breast-ovarian cancer in Icelandic families, and possible relationship to prostatic-cancer. *Am J Hum Genet* 52: 711-717.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (1989). Current Protocols in Molecular Biology. Greene Publishing associates and Wiley-Interscience.
- Ballabio, A. (1993). The rise and fall of positional cloning. *Nature Genet* 3: 277-279.
- Bartkova, J., Lukas, J., Muller, H., Lutzhoft, D., Strauss, M. and Bartek, J. (1994). Cyclin D1 protein expression and function in human breast cancer. *Int J Cancer* 57: 353-61.
- Bellanne-Chantelot, C., Barillot, E., Lacroix, B., Le Paslier, D. and Cohen, D. (1991). A test case for physical mapping of human genome by repetitive sequence fingerprints: construction of a physical map of a 420 kb YAC subcloned into cosmids. *Nucleic Acids Res* 19: 505-10.
- Benedict, W. F., Murphree, A. L., Banerjee, A., Spina, C. A., Sparkes, M. C. and Sparkes, R. S. (1983). Patient with 13 chromosome deletion: evidence that the retinoblastoma gene is a recessive cancer gene. *Science* 219: 973-5.
- Bennion, S. D. and Patterson, J. W. (1984). Keratosis punctata palmaris et plantaris and adenocarcinoma of the colon. A possible familial association of punctate keratoderma and gastrointestinal malignancy. *J Am Acad Dermatol* 10: 587-91.
- Bernstein, J. L., Thompson, W. D., Risch, N. and Holford, T. R. (1992). The genetic epidemiology of second primary breast cancer. *Am J Epidemiol* 136: 937-48.

- Berx, G., Cleton-Jansen, A., Nollet, F., de Leeuw, W., van de Vijver, M., Cornelisse, C. and van Roy, F. (1995). E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. *Embo J* **14**: 6107-15.
- Bianchi, L., Orlandi, A., Iraci, S., Spagnoli, L. G. and Nini, G. (1994). Punctate porokeratotic keratoderma--its occurrence with internal neoplasia. *Clin Exp Dermatol* **19**: 139-41.
- Biggs, J., Hersperger, E., Steeg, P. S., Liotta, L. A. and Shearn, A. (1990). A *Drosophila* gene that is homologous to a mammalian gene associated with tumor metastasis codes for a nucleoside diphosphate kinase. *Cell* **63**: 933-40.
- Birch, J. M. (1990). The Li-Fraumeni cancer family syndrome. *J Pathol* **161**: 1-2.
- Birchmeier, B., Hulsken, J., Behrens, J. (1995). Adherens Junction Proteins in Tumour Progression. In "Cell adhesion and cancer" (I. Hart and N. Hogg, Eds.), Vol. 24, pp. Cancer Surveys. Cold Spring Harbour.
- Bishop, D. T., Cannon, A. L., McLellan, T., Gardner, E. J. and Skolnick, M. H. (1988). Segregation and linkage analysis of nine Utah breast cancer pedigrees. *Genet Epidemiol* **5**: 151-69.
- Black, D. M., Nicolai, H., Borrow, J. and Solomon, E. (1993). A somatic-cell hybrid map of the long arm of human chromosome-17, containing the familial breast-cancer locus (BRCA1). *Am J Hum Genet* **52**: 702-710.
- Blanchet- Bardon, C., Nazzaro, V., Chevrant, B. J., Espie, M., Kerbrat, P. and Le Marec, B. (1987). Hereditary epidermolytic palmoplantar keratoderma associated with breast and ovarian cancer in a large kindred. *Br J Dermatol* **117**: 363-70.
- Bodmer, W., Bishop, T. and Karran, P. (1994). Genetic steps in colorectal cancer. *Nat Genet* **6**: 217-9.
- Borresen, A. L., Andersen, T. I., Eyfjord, J. E., Cornelis, R. S., Thorlacius, S., Borg, A., Johansson, U., Theillet, C., Scherneck, S., Hartman, S., Cornelisse, C. J., Hovig, E. and Devilee P. (1995). Tp53 mutations and breast-cancer prognosis - particularly poor survival rates for cases with mutations in the zinc-binding domains. *Genes Chromosomes Cancer* **14**: 71-75.
- Botstein, D., White, R. L., Skolnick, M. and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* **32**: 314-31.
- Bowcock, A. M., Hall, J. M., Hebert, J. M. and King, M. C. (1990). Exclusion of the retinoblastoma gene and chromosome 13q as the site of a primary lesion for human breast cancer. *Am J Hum Genet* **46**: 12-7.
- Bowcock, A. M., Anderson, L. A., Friedman, L. S., Black, D. M., Osborne, S., Rowell, S. E., Hall, J. M., Solomon, E. and King, M. C. (1993). THRA1 and D17S183 flank an interval of less-than-4 cm for the breast-ovarian cancer gene (BRCA1) on chromosome 17q21. *Am J Hum Genet* **52**: 718-722.
- Boyd, M., Harris, F., McFarlane, R., Davidson, H. R. and Black, D. M. (1995). A human BRCA1 gene knockout. *Nature*. **375**: 541-542.
- Brown, M. A., Jones, K. A., Nicolai, H., Bonjardim, M., Black, D., McFarlane, R., de Jong, P., Quirk, J. P., Lehrach, H. and Solomon, E. (1995). Physical mapping, cloning, and identification of genes within a 500-kb region containing BRCA1. *Proc Natl Acad Sci U S A* **92**: 4362-6.

- Buckle, V.J. and Craig I.W. (1986) In situ hybridisation. in *Human Genetic Diseases: A practical approach*. p85-100. ed. K. E. Davies. IRL Press, Oxford.
- Buckler, A. J., Chang, D. D., Graw, S. L., Brook, J. D., Haber, D. A., Sharp, P. A. and Housman, D. E. (1991). Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc Natl Acad Sci U S A* **88**: 4005-9.
- Buckley, M. F., Sweeney, K. J., Hamilton, J. A., Sini, R. L., Manning, D. L., Nicholson, R. I., deFazio, A., Watts, C. K., Musgrove, E. A. and Sutherland, R. L. (1993). Expression and amplification of cyclin genes in human breast cancer. *Oncogene* **8**: 2127-33.
- Burke, D. T., Carle, G. F. and Olson, M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**: 806-12.
- Cannon, L., Bishop, D. T., Skolnick, M., Hunt, S., Lyon, J. L., Smart, C. R. (1982). Genetic epidemiology of prostate cancer in the Utah Mormon genealogy. *Can Surveys* **1**: 47-69.
- Cantley, L. C., Auger, K. R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. and Soltoff, S. (1991). Oncogenes and signal transduction. *Cell* **64**: 281-302.
- Cavenee, W. K., Dryja, T. P., Phillips, R. A., Benedict, W. F., Godbout, R., Gallie, B. L., Murphree, A. L., Strong, L. C. and White, R. L. (1983). Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature* **305**: 779-84.
- Cavenee, W. K. and White, R. L. (1995). The genetic basis of cancer. *Sci Am* **272**: 72-9.
- Chamberlain, J. S., Boehnke, M., Frank, T. S., Kioussis, S., Xu, J., Guo, S. W., Hauser, E. R., Norum, R. A., Helmbold, E. A., Markel, D. S. et al. (1993). BRCA1 maps proximal to D17S579 on chromosome 17q21 by genetic analysis. *Am J Hum Genet* **52**: 792-8.
- Chen, L. C., Matsumura, K., Deng, G., Kurisu, W., Ljung, B. M., Lerman, M. I., Waldman, F. M. and Smith, H. S. (1994). Deletion of 2 separate regions on chromosome 3p in breast cancers. *Cancer Res* **54**: 3021-3024.
- Chumakov, I. M., Rigault, P., Legall, I., Bellannechantelot, C., Billault, A., Guillou, S., Soularue, P., Guasconi, G., Poullier, E., et al. (1995). A YAC contig map of the human genome. *Nature* **377 (Supp)**: 175-297.
- Claus, E. B., Risch, N. J. and Thompson, W. D. (1990). Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol* **131**: 961-72.
- Claus, E. B., Risch, N. and Thompson, W. D. (1991). Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* **48**: 232-42.
- Cleton-Jansen, J. A., Moerland, E. W., Kuipers, D. N., Callen, D. F., Sutherland, G. R., Hansen, B., Devilee, P. and Cornelisse, C. J. (1994). At least two different regions are involved in allelic imbalance on chromosome arm 16q in breast cancer. *Genes Chromosomes Cancer* **9**: 101-7.
- Cleton-Jansen, A. M., Collins, N., Lakhani, S. R., Weissenbach, J., Devilee, P., Cornelisse, C. J. and Stratton, M. R. (1995). Loss of heterozygosity in sporadic breast tumours at the BRCA2 locus on chromosome 13q12-q13. *Br J Cancer*. **72**: 1241-1244.
- Cohen, D., Chumakov, I. and Weissenbach, J. (1993). A 1st-generation physical map of the human genome. *Nature* **366**: 698-701.
- Coles, C., Condie, A., Chetty, U., Steel, C. M., Evans, H. J. and Prosser, J. (1992). P53 mutations in breast-cancer. *Cancer Res* **52**: 5291-5298.

- Collins, N., McManus, R., Wooster, R., Mangion, J., Seal, S., Lakhani, S. R., Ormiston, W., Daly, P. A., Ford, D., Easton, D. F., *et al.* (1995). Consistent loss of the wild type allele in breast cancers from a family linked to the BRCA2 gene on chromosome 13q12-13. *Oncogene* **10**: 1673-5.
- Cooper, J. A. (1990). Oncogenes and anti-oncogenes. *Curr Opin Cell Biol* **2**: 285-95.
- Cornelis, R. S., Devilee, P., Vanvliet, M., Kuipersdijkshoorn, N., Kersenmaecker, A., Bardoel, A., Khan, P. M. and Cornelisse, C. J. (1993). Allele loss patterns on chromosome-17q in 109 breast carcinomas indicate at least 2 distinct target regions. *Oncogene* **8**: 781-785.
- Cornelis, R. S., Neuhausen, S. L., Johansson, O., Arason, A., Kelsell, D., Ponder, B., Tonin, P., Hamann, U., Lindblom, A., Lalle, P., *et al.* (1995). High allele loss rates at 17q12-q21 in breast and ovarian tumors from BRCA1-linked families. *Genes Chromosomes Cancer*. **13**: 203-210.
- Cottingham, R. J., Idury, R. M. and Schaffer, A. A. (1993). Faster sequential genetic linkage computations. *Am J Hum Genet* **53**: 252-63.
- Coulombe, P. A., Hutton, M. E., Letai, A., Hebert, A., Paller, A. S. and Fuchs, E. (1991). Point mutations in human keratin 14 genes of epidermolysis bullosa simplex patients: genetic and functional analyses. *Cell* **66**: 1301-11.
- Cox, S. A., Vergnaud, G., Milan, J., Moreno, F., Beckmann, J., Kruse, T. and Contu, L. (1994). The EUROGEN map of human chromosome 17. *Eur J Hum Genet* **2**: 236-7.
- Cox, L. S. and Lane, D. P. (1995). Tumour suppressors, kinases and clamps: how p53 regulates the cell cycle in response to DNA damage. *Bioessays* **17**: 501-8.
- Cross, S. H. and Bird, A. P. (1995). CpG islands and genes. *Curr Opin Genet Dev* **5**: 309-14.
- Cross, S. M., Sanchez, C. A., Morgan, C. A., Schimke, M. K., Ramel, S., Idzerda, R. L., Raskind, W. H. and Reid, B. J. (1995). A p53-dependent mouse spindle checkpoint. *Science* **267**: 1353-6.
- Devilee, P., van der. B. M., Kuipers, D. N., Kolluri, R., Khan, P. M., Pearson, P. L. and Cornelisse, C. J. (1989). At least four different chromosomal regions are involved in loss of heterozygosity in human breast carcinoma. *Genomics* **5**: 554-60.
- Devilee, P., Cornelisse, C. J., Kuipers, D. N., Jonker, C. and Pearson, P. L. (1990). Loss of heterozygosity on 17p in human breast carcinomas: defining the smallest common region of deletion. *Cytogenet Cell Genet* **53**: 52-4.
- Devilee, P., Van, V. M., Van, S. P., Dijkshoorn, N. K., Hermans, J., Pearson, P. L. and Cornelisse, C. J. (1991). Allelotype of human breast carcinoma: A second major site for loss of heterozygosity is on chromosome 6q. *Oncogene* **6**: 1705-1711.
- Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J. and Weissenbach, J. (1996). A comprehensive genetic-map of the human genome based on 5,264 microsatellites. *Nature* **380**: 152-154.
- Donehower, L. A. and Bradley, A. (1993). The tumor suppressor p53. *Biochim Biophys Acta* **1155**: 181-205.
- Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., Montgomery, C. A., Butel, J. S. and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumors. *Nature* **356**: 215-221.

- Donis Keller, H., Green, P., Helms, C., Cartinhour, S., Weiffenbach, B., Stephens, K., Keith, T. P., Bowden, D. W., Smith, D. R., *et al.* (1987). A genetic linkage map of the human genome. *Cell* **51**: 319-37.
- Easton, E., Ford, D., Peto, J. (1993a). Inherited susceptibility to breast cancer. In "Breast Cancer" (I. S. Fentiman and J. Taylor-Papadimitriou, Eds.), Vol. 18, pp. 95-113. Cancer Surveys. Cold Spring Harbour Laboratory Press.
- Easton, D. F., Bishop, D. T., Ford, D. and Crockford, G. P. (1993b). Genetic-linkage analysis in familial breast and ovarian-cancer - results from 214 families. *Am J Hum Genet* **52**: 678-701
- Economou, E. P., Bergen, A. W., Warren, A. C. and Antonarakis, S. E. (1990). The polydeoxyadenylate tract of Alu repetitive elements is polymorphic in the human genome. *Proc Natl Acad Sci U S A* **87**: 2951-4.
- Eisinger, F., Stoppa, L. D., Longy, M., Kerangueven, F., Noguchi, T., Bailly, C., Vincent, S. A., Jacquemier, J., Birnbaum, D. and Sobol, H. (1996). Germline mutation at BRCA1 affects the histoprosthetic grade in hereditary breast cancer. *Cancer Res.* **56**: 471-474.
- El-Deiry, W., Tokino, T., Velculescu, V. E., Levy, D. B., Parsons, R., Trent, J. M., Lin, D., Mercer, W. E., Kinzler, K. W. and Vogelstein, B. (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**: 817-25.
- Ellis, A., Field, J. K., Field, E. A., Friedmann, P. S., Fryer, A., Howard, P., Leigh, I. M., Risk, J., Shaw, J. M. and Whittaker, J. (1994). Tylosis associated with carcinoma of the esophagus and oral leukoplakia in a large Liverpool family - a review of 6 generations. *Oral Oncol Eur J Cancer B*
- Elston, R. C., and Lange, K. L. (1975). *Ann. Hum. Genet.* **38**: 341-347.
- Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., Kinzler, K. W. *et al.* (1990). Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science* **247**: 49-56.
- Fearon, E. R. and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell* **61**: 759-67.
- Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**: 6-13.
- Fitzgerald, M. G., MacDonald, D. J., Krainer, M., Hoover, I., O'Neil, E., Unsal, H., Silva, A. S., Finkelstein, D. M., Beer, R. P., *et al.* (1996). Germ-line BRCA1 mutations in Jewish and non-Jewish women with early-onset breast cancer. *New Engl J Med.* **334**: 143-149.
- Ford, D., Easton, D. F., Bishop, D. T., Narod, S. A. and Goldgar, D. E. (1994). Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet* **343**: 692-5.
- Foulds, L. (1957). Tumor progression. *Can. Res* **17**: 355-356.
- Foulkes, W. D., Black, D. M., Stamp, G., Solomon, E. and Trowsdale, J. (1993). Very frequent loss of heterozygosity throughout chromosome 17 in sporadic ovarian carcinoma. *Int J Cancer* **54**: 220-225.
- Friedman, L. S., Szabo, C. I., Ostermeyer, E. A., Dowd, P., Butler, L., Park, T., Lee, M. K., Goode, E. L., Rowell, S. E. and King, M. C. (1995). Novel inherited mutations and variable expressivity of BRCA1 alleles, including the founder mutation 185delAG in Ashkenazi Jewish families. *Am J Hum Genet.* **57**: 1284-1297.

Friend, S. H., Bernard, R., Rogelj, S., Weinberg, R. A., Rapaport, J. M., Albert, D. M. and Dryja, T. P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* **323**: 643-6.

Friend, S., Borresen, A. L., Brody, L., Casey, G., Devilee, P., Gayther, S., Goldgar, D., Murphy, P., Weber, B. L. and Wiseman, R. (1995). Breast cancer information on the web. *Nat Genet* **11**: 238-9.

Fuchs, E. and Coulombe, P. A. (1992). Of mice and men - genetic skin diseases of keratin. *Cell* **69**: 899-902.

Futreal, P. A., Soderkvist, P., Marks, J. R., Iglehart, J. D., Cochran, C., Barrett, J. C. and Wiseman, R. W. (1992). Detection of frequent allelic loss on proximal chromosome 17q in sporadic breast carcinoma using microsatellite length polymorphisms. *Cancer Res* **52**: 2624-7.

Futreal, P. A., Liu, Q., Shattuck, E. D., Cochran, C., Harshman, K., Tavtigian, S., Bennet, L. M., Haugen, S. A., Swensen, J., Miki, Y., Eddington, K., *et al.* (1994). BRCA1 mutations in primary breast and ovarian carcinomas. *Science* **266**: 120-122.

Garrod, D. R. (1995). Desmosomes and cancer. In "Cell Adhesion and Cancer" (I. Hart and N. Hogg, Eds.), Vol. 24, pp. 97-111. Cold Spring Harbour Laboratory Press.

Gayther, S. A., Warren, W., Mazoyer, S., Russell, P. A., Harrington, P. A., Chiano, M., Seal, S., Hamoudi, R., Van, R. E., Dunning, A. M., Love, R., *et al.* (1995). Germline mutations of the BRCA1 gene in breast and ovarian cancer families provide evidence for a genotype-phenotype correlation. *Nat Genet.* **11**: 428-433.

Gilles, A. M., Presecan, E., Vonica, A. and Lascu, I. (1991). Nucleoside diphosphate kinase from human erythrocytes. Structural characterization of the two polypeptide chains responsible for heterogeneity of the hexameric enzyme. *J Biol Chem* **266**: 8784-9.

Goldgar, D. E., Fields, P., Lewis, C. M., Tran, T. D., Cannon-Albright, L. A. and Ward, J. H. (1994). A large kindred with 17q-linked breast and ovarian-cancer - genetic, phenotypic, and genealogical analysis. *J Natl Cancer Inst* **86**: 200-209.

Goodrich, D. W., Wang, N. P., Qian, Y. W., Lee, E. Y. and Lee, W. H. (1991). The retinoblastoma gene product regulates progression through the G1 phase of the cell cycle. *Cell* **67**: 293-302.

Gowen, L. C., Johnson, B. L., Latour, A. M., Sulik, K. K. and Koller, B. H. (1996). Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nat Genet* **12**: 191-4.

Gray, I. C., Phillips, S. M. A., Lee, S. J., Neoptolemos, J. P., Weissenbach, J. and Spurr, N. K. (1995). Loss of the chromosomal region 10q23-25 in prostate-cancer. *Cancer Res* **55**: 4800-4803.

Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., Robertson, M. *et al.* (1991). Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* **66**: 589-600.

Gudas, J. M., Nguyen, H., Li, T. and Cowan, K. H. (1995). Hormone-dependent regulation of BRCA1 in human breast cancer cells. *Cancer Res* **55**: 4561-5.

Gudmundsson, J., Johannesdottir, G., Bergthorsson, J. T., Arason, A., Ingvarsson, S., Egilsson, V. and Barkardottir, R. B. (1995). Different tumor types from BRCA2 carriers show wild-type chromosome deletions on 13q12-q13. *Cancer Res* **55**: 4830-2.

- Gudmundsson, J., Johannesdottir, G., Arason, A., Bergthorsson, J. T., Ingvarsson, S., Egilsson, V., Barkardottir, R. B. (1996). Frequent occurrence of BRCA2 linkage in Icelandic breast cancer families and segregation of a common BRCA2 haplotype. *Am. J. Hum. Genet* **58**: 749-756.
- Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y. *et al.* (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**: 234-8.
- Gusterson, B. A., Stratton, M., Anbazhagan, R. The importance of human breast development in mammary carcinogenesis. In Li SA, Gustafsson, J-A, Nandi, S., Sekely, I. I., eds "Hormonal Carcinogenesis: proceedings of the second international symposium. New York; Springer, 1996: 115-119.
- Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M. and Weissenbach, J. (1994). The 1993-94 Genethon human genetic linkage map. *Nat Genet*. **7**. 246-339.
- Hall, J. M., Lee, M. K., Newman, B., Morrow, J. E., Anderson, L. A., Huey, B. and King, M. C. (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* **250**: 1684-9.
- Hamada, H., Petrino, M. G., Kakunaga, T., Seidman, M. and Stollar, B. D. (1984). Characterization of genomic poly(dT-dG).poly(dC-dA) sequences: structure, organization, and conformation. *Mol Cell Biol* **4**: 2610-21.
- Han, H. J., Fujiwara, T., Shin, S. and Nakamura, Y. (1993). Dinucleotide repeat polymorphism in the 3' non-coding region of the FLT1 gene. *Hum Mol Genet* **2**: ???
- Hansen, M. F. and Cavenee, W. K. (1987). Genetics of cancer predisposition. *Cancer Res* **47**: 5518-27.
- Hanssen, A. M. and Fryns, J. P. (1995). Cowden syndrome. *J Med Genet* **32**: 117-9.
- Hartwell, L. H. and Kastan, M. B. (1994). Cell-cycle control and cancer. *Science* **266**: 1821-1828.
- Hawthorn, L. A. and Cowell, J. K. (1996). Regional assignment of EST sequences on human chromosome 13. *Cytogenet Cell Genet* **72**: 72-7.
- Hearne, C. M., Ghosh, S. and Todd, J. A. (1992). Microsatellites for linkage analysis of genetic traits. *Trends Genet* **8**: 288-94.
- Hennies, H. C., Hagedorn, M. and Reis, A. (1995a). Palmoplantar keratoderma in association with carcinoma of the esophagus maps to chromosome 17q distal to the keratin gene-cluster. *Genomics* **29**: 537-540.
- Hennies, H. C., Kuster, W., Mischke, D. and Reis, A. (1995b). Localization of a locus for the striated form of palmoplantar keratoderma to chromosome 18q near the desmosomal cadherin gene-cluster. *Hum Mol Genet* **4**: 1015-1020.
- Herman, J. G., Latif, F., Weng, Y., Lerman, M. I., Zbar, B., Liu, S., Samid, D., Duan, D. S., Gnarr, J. R., Linehan, W. M. *et al.* (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A* **91**: 9700-4.
- Hofstra, R. M., Landsvater, R. M., Ceccherini, I., Stulp, R. P., Stelwagen, T., Luo, Y., Pasini, B., Hoppener, J. W., van, A. H., Romeo, G. *et al.* (1994). A mutation in the RET proto-oncogene associated with multiple endocrine neoplasia type 2B and sporadic medullary thyroid carcinoma. *Nature* **367**: 375-6.

- Hollstein, M., Sidransky, D., Vogelstein, B. and Harris, C. C. (1991). p53 mutations in human cancers. *Science* **253**: 49-53.
- Holt, J. T., Thompson, M. E., Szabo, C., Robinson, B. C., Arteaga, C. L., King, M. C. and Jensen, R. A. (1996). Growth retardation and tumour inhibition by BRCA1. *Nat Genet* **12**: 298-302.
- Howell-Evans, W., McConnell, R. B., Clarke, C. A., Sheppard, P. M. (1958). Carcinoma of the oesophagus with keratosis palmaris et plantaris (tylosis). *Quart. J. Med.* **28**: 413-429.
- Hussussian, C. J., Struewing, J. P., Goldstein, A. M., Higgins, P. A., Ally, D. S., Sheahan, M. D., Clark, W. J., Tucker, M. A. and Dracopoli, N. C. (1994). Germline p16 mutations in familial melanoma. *Nat Genet* **8**: 15-21.
- Ikeda, S., Welsh, E. A., Peluso, A. M., Leyden, W., Duvic, M., Woodley, D. T. and Epstein, E. H. (1994). Localization of the gene whose mutations underlie Hailey-Hailey disease to chromosome 3q. *Hum Mol Genet* **3**: 1147-1150.
- Iselius, L., Slack, J., Littler, M. and Morton, N. E. (1991). Genetic epidemiology of breast cancer in Britain. *Ann Hum Genet* **55**: 151-159.
- Jacks, T., Remington, L., Williams, B. O., Schmitt, E. M., Halachmi, S., Bronson, R. T. and Weinberg, R. A. (1994). Tumor spectrum analysis in p53-mutant mice. *Curr Biol* **4**: 1-7.
- Jacobs, I. J., Smith, S. A., Wiseman, R. W., Futreal, P. A., Harrington, T., Osborne, R. J., Leech, V., Molyneux, A., Berchuck, A., Ponder, B. and Bast, R. C. (1993). A deletion unit on chromosome-17q in epithelial ovarian-tumors distal to the familial breast ovarian-cancer locus. *Cancer Res* **53**: 1218-1221.
- Jacquemier, J., Eisinger, F., Birnbaum, D. and Sobol, H. (1995). Histoprognostic grade in BRCA1-associated breast cancer. *Lancet* **345**: 1903.
- Jeffreys, A. J., Wilson, V. and Thein, S. L. (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature* **314**: 67-73.
- Jensen, R. A., Thompson, M. E., Jetton, T. L., Szabo, C. I., van der Meer, R., Helou, B., Tronick, S. R., Page, D. L., King, M. C. and Holt, J. T. (1996). BRCA1 is secreted and exhibits properties of a granin. *Nat Genet* **12**: 303-8.
- Jinno, Y., Yun, K., Nishiwaki, K., Kubota, T., Ogawa, O., Reeve, A. E. and Niikawa, N. (1994). Mosaic and polymorphic imprinting of the WT1 gene in humans. *Nat Genet* **6**: 305-9.
- Jones, M. H. and Nakamura, Y. (1992). Detection of loss of heterozygosity at the human tp53 locus using a dinucleotide repeat polymorphism. *Genes Chromosomes Cancer* **5**: 89-90.
- Kallioniemi, A., Kallioniemi, O. P., Piper, J., Tanner, M., Stokke, T., Chen, L., Smith, H. S., Pinkel, D., Gray, J. W. and Waldman, F. M. (1994). Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc Natl Acad Sci U S A* **91**: 2156-60.
- Kamb, A., Gruis, N. A., Weaverfeldhaus, J., Liu, Q. Y., Harshman, K., Tavtigian, S. V., Stockert, E., Day, R. S., Johnson, B. E. and Skolnick, M. H. (1994). A cell-cycle regulator potentially involved in genesis of many tumor types. *Science* **264**: 436-440.
- Kampert, J. B., Whittemore, A. S. and Paffenbarger, R. J. (1988). Combined effect of childbearing, menstrual events, and body size on age-specific breast cancer risk. *Am J Epidemiol* **128**: 962-79.

- Karunagaran, D., Tzahar, E., Beerli, R. R., Chen, X., Graus, P. D., Ratzkin, B. J., Seger, R., Hynes, N. E. and Yarden, Y. (1996). ErbB-2 is a common auxiliary subunit of NDF and EGF receptors: Implications for breast cancer. *Embo J.* **15**: 254-264.
- Kastan, M. B., Onyekwere, O., Sidransky, D., Vogelstein, B. and Craig, R. W. (1991). Participation of p53 protein in the cellular response to DNA damage. *Cancer Res.* **51**: 6304-6311.
- Kelsell, D. P., Black, D. M., Bishop, D. T. and Spurr, N. K. (1993a). Genetic analysis of the BRCA1 region in a large breast/ovarian family: refinement of the minimal region containing BRCA1. *Hum Mol Genet* **2**: 1823-8.
- Kelsell, D. P., Black, D. M., Solomon, E. and Spurr, N. K. (1993b). Localization of a 2nd nm23 gene, NME2, to chromosome-17q21-q22. *Genomics* **17**: 522-524.
- Kelsell, D. P., Rooke, L., Warne, D., Bouzyk, M., Cullin, L., Cox, S., West, L., Povey, S. and Spurr, N. K. (1995a). Development of a panel of monochromosomal somatic-cell hybrids for rapid gene-mapping. *Ann Hum Genet* **59**: 233-241.
- Kelsell, D. P., Stevens, H. P., Ratnavel, R., Bryant, S. P., Bishop, D. T., Leigh, I. M. and Spurr, N. K. (1995b). Genetic-linkage studies in non-epidermolytic palmoplantar keratoderma - evidence for heterogeneity. *Hum Mol Genet* **4**: 1021-1025.
- Kelsell, D. P., Spurr, N. K., Barnes, D. M., Gusterson, B., Bishop, D. T. (1996a). Combined loss of BRCA1/ BRCA2 in grade 3 breast carcinomas. *Lancet.* **347**: 1554-1555.
- Kelsell, D. P., Risk, J. M., Leigh, I. M., Stevens, H. P., Ellis, A., Hennies, H. C., Reis, A., Weissenbach, J., Bishop, D. T., Spurr, N. K., Field, J. K. (1996b). Close mapping of the focal non-epidermolytic palmoplantar keratoderma (PPK) locus associated with oesophageal cancer (TOC). *Hum. Mol. Genet.* **5**: 857-860.
- Kerangueven, F., Essioux, L., Dib, A., Noguchi, T., Allione, F., Geneix, J., Longy, M., Lidereau, R., Eisinger, F. and Pebusque, M. J. (1995). Loss of heterozygosity and linkage analysis in breast-carcinoma - indication for a putative 3rd susceptibility gene on the short arm of chromosome-8. *Oncogene* **10**: 1023-1026.
- Kibar, Z., Der Kaloustian, V. M., Brais, B., Hani, V., Fraser, F. C., Rouleau, G. A. (1996). The gene responsible for Clouston hidrotic ectodermal dysplasia maps to the pericentric region of chromosome 13q. *Hum. Mol. Genet.* **5**: 543-547.
- Kimonis, V., DiGiovanna, J. J., Yang, J. M., Doyle, S. Z., Bale, S. J. and Compton, J. G. (1994). A mutation in the V1 end domain of keratin 1 in non-epidermolytic palmar-plantar keratoderma. *J Invest Dermatol* **103**: 764-9.
- King, M. C. (1992). Breast cancer genes: how many, where and who are they?. *Nat Genet* **2**: 89-90.
- Knudson, A. J. (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* **68**: 820-3.
- Knudson, A. J., Meadows, A. T., Nichols, W. W. and Hill, R. (1976). Chromosomal deletion and retinoblastoma. *N Engl J Med* **295**: 1120-3.
- Kooy, R. F., Wijngaard, A., Verlind, E., Vergnaud, G., Scheffer, H. and Buys, C. H. (1994). The EUROGEN map of human chromosome 13. *Eur J Hum Genet* **2**: 228-9.
- Kovach, J. S., Hartmann, A., Blaszyk, H., Cunningham, J., Schaid, D. and Sommer, S. S. (1996). Mutation detection by highly sensitive methods indicates that p53 gene mutations in breast cancer can have important prognostic value. *Proc Natl Acad Sci U S A.* **93**: 1093-1096.

- Kranenburg, O., van der Eb, A. and Zantema, A. (1996). Cyclin D1 is an essential mediator of apoptotic neuronal cell death. *Embo J* **15**: 46-54.
- Krauter, K., Montgomery, K., Yoon, S-Y., LeBlanc-Straceski, J., Renault, B., Marondel, I., Herdman, V., *et al.* (1995). A second generation YAC contig map of human chromosome 12. *Nature*. **377** **supp**: 321-333.
- Kuerbitz, S. J., Plunkett, B. S., Walsh, W. V. and Kastan, M. B. (1992). Wild-type p53 is a cell cycle checkpoint determinant following irradiation. *Proc Natl Acad Sci U S A* **89**: 7491-5.
- Lancaster, J. M., Wooster, R., Mangion, J., Phelan, C. M., Cochran, C., Gumbs, C., Seal, S., Barfoot, R., Collins, N., *et al.* (1996). BRCA2 mutations in primary breast and ovarian cancers. *Nat. Genet.* **13**: 238-240.
- Lane, E. B., Rugg, E. L., Navsaria, H., Leigh, I. M., Heagerty, A. H., Ishida, Y. A. and Eady, R. A. (1992). A mutation in the conserved helix termination peptide of keratin 5 in hereditary skin blistering. *Nature* **356**: 244-6.
- Lane, E. B. (1994). Keratin diseases. *Curr Opin Genet Dev* **4**: 412-8.
- Langbein, L., Heid, H. W., Moll, I. and Franke, W. W. (1993). Molecular characterization of the body site-specific human epidermal cytokeratin 9: cDNA cloning, amino acid sequence, and tissue specificity of gene expression. *Differentiation* **55**: 57-71.
- Langston, A. A., Malone, K. E., Thompson, J. D., Daling, J. R. and Ostrander, E. A. (1996). BRCA1 mutations in a population-based sample of young women with breast cancer. *N Engl J Med* **334**: 137-42.
- Lasko, D., Cavenee, W. and Nordenskjold, M. (1991). Loss of constitutional heterozygosity in human cancer. *Annu Rev Genet* **25**: 281-314.
- Lathrop, G. M. and Lalouel, J. M. (1984). Easy calculations of lod scores and genetic risks on small computers. *Am J Hum Genet* **36**: 460-5.
- Lathrop, G. M., Lalouel, J. M. and White, R. L. (1986). Construction of human linkage maps: likelihood calculations for multilocus linkage analysis. *Genet Epidemiol* **3**: 39-52.
- Lathrop, G. M. and Lalouel, J. M. (1988). Efficient computations in multilocus linkage analysis. *Am J Hum Genet* **42**: 498-505.
- Latif, F., Tory, K., Gnarr, J., Yao, M., Duh, F. M., Orcutt, M. L., Stackhouse, T., Kuzmin, I., Modi, W., Geil, L., Schmidt, L., Zhou, F. W., Li, H., *et al.* (1993). Identification of the vonhippel-lindau disease tumor-suppressor gene. *Science* **260**: 1317-1320.
- Lee, E. Y., To, H., Shew, J. Y., Bookstein, R., Scully, P. and Lee, W. H. (1988). Inactivation of the retinoblastoma susceptibility gene in human breast cancers. *Science* **241**: 218-21.
- Lessin, S. R., Huebner, K., Isobe, M., Croce, C. M. and Steinert, P. M. (1988). Chromosomal mapping of human keratin genes: evidence of non-linkage. *J Invest Dermatol* **91**: 572-8.
- Lind, L., Lundstrom, A., Hofer, P. A. and Holmgren, G. (1994). The gene for diffuse palmoplantar keratoderma of the type found in northern Sweden is localized to chromosome 12q11-q13. *Hum Mol Genet* **3**: 1789-93.
- Lindahl, T. (1994). DNA repair. DNA surveillance defect in cancer cells. *Curr Biol* **4**: 249-51.
- Lindblom, A., Skoog, L., Rotstein, S., Werelius, B., Larsson, C. and Nordenskjold, M. (1993). Loss of heterozygosity in familial breast carcinomas. *Cancer Res* **53**: 4356-4361.

- Livingstone, L. R., White, A., Sprouse, J., Livanos, E., Jacks, T. and Tlsty, T. D. (1992). Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. *Cell* **70**: 923-35.
- Lobaccaro, J. M., Lumbroso, S., Belon, C., Galtier, D. F., Bringer, J., Lesimple, T., Namer, M., Cutuli, B. F., Pujol, H. and Sultan, C. (1993). Androgen receptor gene mutation in male breast cancer. *Hum Mol Genet* **2**: 1799-802.
- Loeb, L. A. (1991). Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res* **51**: 3075-9.
- Lucker, G. P., van der Kerkhof, K. P. and Steijlen, P. M. (1994). The hereditary palmoplantar keratoses: an updated review and classification. *Br J Dermatol* **131**: 1-14.
- Luu, T. V., Labrie, C., Simard, J., Lachance, Y., Zhao, H. F., Couet, J., Leblanc, G. and Labrie, F. (1990). Structure of two in tandem human 17 beta-hydroxysteroid dehydrogenase genes. *Mol Endocrinol* **4**: 268-75.
- Lynch, H. T., Harris, R. E., Guirgis, H. A., Maloney, K., Carmody, L. L. and Lynch, J. F. (1978). Familial association of breast/ovarian carcinoma. *Cancer* **41**: 1543-9.
- Mackay, J., Steel, C. M., Elder, P. A., Forrest, A. P. and Evans, H. J. (1988). Allele loss on short arm of chromosome 17 in breast cancers. *Lancet* **2**: 1384-5.
- MacPhee, D. G. (1995). Mismatch repair, somatic mutations, and the origins of cancer. *Cancer Res* **55**: 5489-92.
- Maestrini, E., Monaco, A. P., McGrath, J. A., Ishida-Yamamoto, A., Camisa, C., Hovnanian, A., Weeks, D. E., Lathrop, M., Uitto, J., Christiano, A. M. (1996). A molecular defect in lorcin, the major component of the cornified cell envelope, underlies Vohwinkel's syndrome. *Nat. Genet.* **13**: 70-77.
- Malkin, D. and Friend, S. H. (1993). Germ line p53 mutations in a familial syndrome of breast-cancer, sarcomas, and other neoplasms. *Science* **259**: 878-878.
- Mansfield, D. C., Brown, A. F., Green, D. K., Carothers, A. D., Morris, S. W., Evans, H. J. and Wright, A. F. (1994). Automation of genetic-linkage analysis using fluorescent microsatellite markers. *Genomics* **24**: 225-233.
- Marger, R. S. and Marger, D. (1993). Carcinoma of the esophagus and tylosis - a lethal genetic combination. *Cancer* **72**: 17-19.
- Marquis, S. T., Rajan, J. V., Wynshaw, B. A., Xu, J., Yin, G. Y., Abel, K. J., Weber, B. L. and Chodosh, L. A. (1995). The developmental pattern of Brca1 expression implies a role in differentiation of the breast and other tissues. *Nat Genet.* **11**: 17-26.
- Matsushime, H., Quelle, D. E., Shurtleff, S. A., Shibuya, M., Sherr, C. J. and Kato, J. Y. (1994). D-type cyclin-dependent kinase activity in mammalian cells. *Mol Cell Biol* **14**: 2066-76.
- Mazoyer, S., Gayther, S. A., Nagai, M. A., Smith, S. A., Dunning, A., van Rensburg, E., Albertsen, H., White, R. and Ponder, B. A. (1995). A gene (DLG2) located at 17q12-q21 encodes a new homologue of the Drosophila tumor suppressor dIc-A. *Genomics* **28**: 25-31.
- McLean, W. H., Eady, R. A., Dopping-Hepenstahl, P., McMillan, J. R., Leigh, I. M., Navsaria, H. A., Higgins, C., Harper, J. I., Paige, D. G., Morley, S. M. *et al.* (1994). Mutations in the rod 1A domain of keratins 1 and 10 in bullous congenital ichthyosiform erythroderma (BCIE). *J Invest Dermatol* **102**: 24-30.

- McLean, W. H. and Lane, E. B. (1995). Intermediate filaments in disease. *Curr Opin Cell Biol* 7: 118-25.
- McLean, W. H. I., Rugg, E. L., Lunny, D. P., Morley, S. M., Lane, E. B., Swensson, O., Dopping-Hepenstal, P. J. C., Griffiths, W. A. D., Eady, R. A. J. and Higgins, C. (1995). Keratin16 and keratin17 mutations cause pachyonychia-congenita. *Nature Genet* 9: 273-278.
- McNeill, J. M., Reed, M. J., Beranek, P. A., Bonney, R. C., Ghilchik, M. W., Robinson, D. J. and James, V. H. (1986). A comparison of the in vivo uptake and metabolism of 3H-oestrone and 3H-oestradiol by normal breast and breast tumour tissues in post-menopausal women. *Int J Cancer* 38: 193-6.
- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennet, L. M., Ding, W., Bell, R., *et al.* (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266: 66-71.
- Miki, Y., Katagiri, T., Kasumi, F., Yoshimoto, T., Nakamura, Y. (1996). Mutation analyses in the BRCA2 gene in primary breast cancers. *Nat. Genet.* 13: 245-247.
- Mitelman, F. and Levan, G. (1981). Clustering of aberrations to specific chromosomes in human neoplasms. IV. A survey of 1,871 cases. *Hereditas* 95: 79-139.
- Moll, I., Heid, H. W., Franke, W. W. and Moll, R. (1988). Patterns of expression of trichocytic and epithelial cytokeratins in mammalian tissues. III. Hair and nail formation during human fetal development. *Differentiation* 39: 167-84.
- Mori, T., Aoki, T., Matsubara, T., Iida, F., Du, X., Nishihira, T., Mori, S. and Nakamura, Y. (1994). Frequent loss of heterozygosity in the region including BRCA1 on chromosome 17q in squamous cell carcinomas of the esophagus. *Cancer Res* 54: 1638-40.
- Morton, N. (1955). Sequential tests for the detection of linkage. *Am. J. Hum. Genet.* 7: 277-318.
- Moscattello, D. K., Holgado, M. M., Godwin, A. K., Ramirez, G., Gunn, G., Zoltick, P. W., Biegel, J. A., Hayes, R. L. and Wong, A. J. (1995). Frequent expression of a mutant epidermal growth factor receptor in multiple human tumors. *Cancer Res* 55: 5536-9.
- Mulligan, L. M., Kwok, J., Healey, C. S., Elsdon, M. J., Eng, C., Gardner, E., Love, D. R., Mole, S. E., Moore, J. K., Papi, L., Ponder, M. A., Telenius, H., Tunnacliffe, A. and Ponder, B. (1993). Germ-line mutations of the ret protooncogene in multiple endocrine neoplasia type-2a. *Nature* 363: 458-460.
- Nagatake, M., Osada, H., Kondo, M., Uchida, K., Nishio, M., Shimokata, K., Takahashi, T., Takahashi, T. (1996). Aberrant hypermethylation at the bcl-2 locus at 18q21 in human lung cancers. *Cancer Res.* 56: 1886-1891.
- Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. *et al.* (1987). Variable number of tandem repeat (VNTR) markers for human gene mapping. *Science* 235: 1616-22.
- Narod, S. A., Feunteun, J., Lynch, H. T., Watson, P., Conway, T., Lynch, J. and Lenoir, G. M. (1991). Familial breast-ovarian cancer locus on chromosome 17q12-q23. *Lancet* 338: 82-3.
- Narod, S. A., Ford, D., Devilee, P., Barkardottir, R. B., Lynch, H. T., Smith, S. A., Ponder, B., Weber, B. L., Garber, J. E., Birch, J. M., Cornelis, R. S., Kelsell, D. P. *et al.* (1995). An evaluation of genetic heterogeneity in 145 breast-ovarian cancer families. *Am J Hum Genet.* 56: 254-264.

Nelen, M. R., Padberg, G. W., Peeters, E. A. J., Lin, A. Y., van den Helm, B., Frants, R. R., Coulon, V., Goldstein, A.M., van Reen, M. M. M., Easton, D. F., *et al.* (1996). Localisation of the gene for Cowden disease to chromosome 10q22-23. *Nat. Genet.* 13: 114-116.

Neuhausen, S. L., Mazoyer, S., Friedman, L., Stratton, M., Offit, K., Caligo, A., Tomlinson, G., Cannon-Allbright, L., Bishop, T., Kelsell, D., Solomon, E., *et al.* (1996a). Haplotype and phenotype analysis of six recurrent BRCA1 mutations in 61 families: Results of an international study. *Am J Hum Genet.* 58: 271-280.

Neuhausen, S., Gilewski, T., Norton, L., Tran, T., McGuire, P., Swensen, J., Hampel, H., Borgen, P., Brown, K., Skolnick, M., Shattuck-Eidens, D., Jhanwar, S., Goldgar, D., Offit, K. (1996b). Recurrent BRCA2 6174delT mutations in Ashkenazi Jewish women affected by breast cancer. *Nat. Genet.* 13: 126-128.

Newman, B., Austin, M. A., Lee, M. and King, M. C. (1988). Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci U S A* 85: 3044-8.

NIH/CEPH collaborative Mapping Group. (1992). A comprehensive genetic-linkage map of the human genome. *Science* 258: 148-162.

Normand, T., Narod, S., Labrie, F. and Simard, J. (1993). Detection of polymorphisms in the estradiol 17 beta-hydroxysteroid dehydrogenase II gene at the EDH17B2 locus on 17q11-q21. *Hum Mol Genet* 2: 479-83.

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science* 194: 23-8.

Oka, H., Shiozaki, H., Kobayashi, K., Inoue, M., Tahara, H., Kobayashi, T., Takatsuka, Y., Matsuyoshi, N., Hirano, S., Takeichi, M., *et al.* (1993). Expression of E-cadherin cell adhesion molecules in human breast cancer tissues and its relationship to metastasis. *Cancer Res* 53: 1696-701.

Oliner, J. D., Kinzler, K. W., Meltzer, P. S., George, D. L., Vogelstein, B. (1992). Amplification of a gene encoding a p53-associated protein in human sarcomas. *Nature.* 358: 80-83.

Osborne, R. J., Merlo, G. R., Mitsudomi, T., Venesio, T., Liscia, D. S., Cappa, A. P., Chiba, I., Takahashi, T., Nau, M. M., Callahan, R., *et al.* (1991). Mutations in the p53 gene in primary human breast cancers. *Cancer Res* 51: 6194-8.

Ott, J. (1992). Analysis of human genetic linkage. Baltimore and London: John Hopkins University Press.

Parfitt, E., Burge, S., Craddock, N., Roberts, E., McLean, W. H., Weissenbach, J., McGuffin, P. and Owen, M. (1994). The gene for Darier's disease maps between D12S78 and D12S79. *Hum Mol Genet* 3: 35-8.

Petrakis N. L., Ernster, V., King, M. C. K. (1982). Breast cancer. In "Cancer epidemiology and prevention" (WB Saunders, Ed.), Vol. pp. 855-870. Philadelphia.

Phelan, C. M., Lancaster, J. M., Tonin, P., Gumbs, C., Cochran, C., Carter, R., Ghadirian, P., Perret, C., Moslehi, R., Dion, F., Faucher, M-C., Dole, K., *et al.* (1996). Mutation analysis of the BRCA2 gene in 49 site-specific breast cancer families. *Nat. Genet.* 13: 120-122.

Phillips, R. L. (1975). Role of life-style and dietary habits in risk of cancer among seventh-day adventists. *Cancer Res*

Phipps, R. F. and Perry, P. M. (1989). Familial breast cancer and the association with colonic carcinoma. *Eur J Surg Oncol* 15: 109-11.

- Ponder, B. A. and Wilkinson, M. M. (1986). Direct examination of the clonality of carcinogen-induced colonic epithelial dysplasia in chimeric mice. *J Natl Cancer Inst* 77: 967-76.
- Prakash, O. and Yunis, J. J. (1984). High resolution chromosomes of the t(9;22) positive leukemias. *Cancer Genet Cytogenet* 11: 361-7.
- Prosser, J., Thompson, A. M., Cranston, G. and Evans, H. J. (1990). Evidence that p53 behaves as a tumour suppressor gene in sporadic breast tumours. *Oncogene* 5: 1573-9.
- Radford, D. M., Fair, K., Thompson, A. M., Ritter, J. H., Holt, M., Steinbrueck, T., Wallace, M., Wells, S. J. and Donis, K. H. (1993). Allelic loss on a chromosome 17 in ductal carcinoma in situ of the breast. *Cancer Res* 53: 2947-9.
- Reed, P. W., Davies, J. L., Copeman, J. B., Bennett, S. T., Palmer, S. M., Pritchard, L. E., Gough, S. C. L., Kawaguchi, Y., Cordell, H. J. and Balfour, K. M. (1994). Chromosome-specific microsatellite sets for fluorescence-based, semiautomated genome mapping. *Nature Genet* 7: 390-395.
- Reis, A., Hennies, H. C., Langbein, L., Digweed, M., Mischke, D., Drechsler, M., Schrock, E., Royer, P. B., Franke, W. W., Sperling, K. et al. (1994). Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK). *Nat Genet* 6: 174-9.
- Riordan, J. R., Rommens, J. M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J. L. et al. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 245: 1066-73.
- Risk, J. M., Field, E. A., Field, J. K., Whittaker, J., Fryer, A., Ellis, A., Shaw, J. M., Friedmann, P. S., Bishop, D. T. and Bodmer, J. (1994). Tylosis esophageal cancer mapped. *Nature Genet* 8: 319-321.
- Roberts, L. (1993). Genetics - zeroing in on a breast-cancer susceptibility gene. *Science* 259: 622.
- Rodgers, C. S., Hill, S. M. and Hulten, M. A. (1984). Cytogenetic analysis in human breast carcinoma. I. Nine cases in the diploid range investigated using direct preparations. *Cancer Genet Cytogenet* 13: 95-119.
- Rosenberg, M., RayChaudhury, A., Shows, T. B., Le, B. M. and Fuchs, E. (1988). A group of type I keratin genes on human chromosome 17: characterization and expression. *Mol Cell Biol* 8: 722-36.
- Rosengard, A. M., Krutzsch, H. C., Shearn, A., Biggs, J. R., Barker, E., Margulies, I. M., King, C. R., Liotta, L. A. and Steeg, P. S. (1989). Reduced Nm23/Awd protein in tumour metastasis and aberrant Drosophila development. *Nature* 342: 177-80.
- Rosnet, O., Stephenson, D., Mattei, M. G., Marchetto, S., Shibuya, M., Chapman, V. M. and Birnbaum, D. (1993). Close physical linkage of the FLT1 and FLT3 genes on chromosome 13 in man and chromosome 5 in mouse. *Oncogene* 8: 173-9.
- Royle, N. J., Clarkson, R. E., Wong, Z. and Jeffreys, A. J. (1988). Clustering of hypervariable minisatellites in the proterminal regions of human autosomes. *Genomics* 3: 352-60.
- Rubinfeld, B., Souza, B., Albert, I., Muller, O., Chamberlain, S. H., Masiarz, F. R., Munemitsu, S. and Polakis, P. (1993). Association of the APC gene product with beta-catenin. *Science* 262: 1731-4.

- Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-91.
- Saito, H., Inazawa, J., Saito, S., Kasumi, F., Koi, S., Sagae, S., Kudo, R., Saito, J., Noda, K. and Nakamura, Y. (1993). Detailed deletion mapping of chromosome 17q in ovarian and breast cancers - 2-cm region on 17q21.3 often and commonly deleted in tumors. *Cancer Res* **53**: 3382-3385.
- Sambrook, J., Fritsch, E. F., Maniatis, T. (1989). *Molecular Cloning: a laboratory manual*. 3 vols. Cold Spring Harbor Laboratory Press.
- Sanger, F., Nicklen, S. and Coulson, A. R. (1992). DNA sequencing with chain-terminating inhibitors. 1977. *Biotechnology* **24**: 104-8.
- Sasaki, M., Okamoto, M., Sato, C., Sugio, K., Soejima, J., Iwama, T., Ikeuchi, T., Tonomura, A., Miyaki, M. and Sasazuki, T. (1989). Loss of constitutional heterozygosity in colorectal tumors from patients with familial polyposis coli and those with nonpolyposis colorectal carcinoma. *Cancer Res* **49**: 4402-6.
- Sato, T., Saito, H., Swensen, J., Olifant, A., Wood, C., Danner, D., Sakamoto, T., Takita, K., Kasumi, F., Miki, Y., Skolnick, M. and Nakamura, Y. (1992). The human prohibitin gene located on chromosome 17q21 is mutated in sporadic breast cancer. *Cancer Res* **52**: 1643-1646.
- Savitsky, K., Bar, S. A., Gilad, S., Rotman, G., Ziv, Y., Vanagaite, L., Tagle, D. A., Smith, S., Uziel, T., Sfez, S., *et al.* (1995). A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* **268**: 1749-53.
- Schaffer, A. A., Gupta, S. K., Shriram, K. and Cottingham, R. J. (1994). Avoiding recomputation in linkage analysis. *Hum Hered* **44**: 225-37.
- Schildkraut, J. M., Risch, N. and Thompson, W. D. (1989). Evaluating genetic association among ovarian, breast, and endometrial cancer: evidence for a breast/ovarian cancer relationship. *Am J Hum Genet* **45**: 521-9.
- Schlessinger, D. (1990). Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Genet* **6**: 255-8.
- Schott, D. R., Chang, J. N., Deng, G., Kurisu, W., Kuo, W. L., Gray, J. and Smith, H. S. (1994). A candidate tumor suppressor gene in human breast cancers. *Cancer Res* **54**: 1393-6.
- Schutte, M., da, C. L., Hahn, S. A., Moskaluk, C., Hoque, A. T., Rozenblum, E., Weinstein, C. L., Bittner, M., Meltzer, P. S., Trent, J. M. *et al.* (1995a). Identification by representational difference analysis of a homozygous deletion in pancreatic carcinoma that lies within the BRCA2 region. *Proc Natl Acad Sci U S A* **92**: 5950-4.
- Sellers, T. A., Kushi, L. H., Potter, J. D., Kaye, S. A., Nelson, C. L., McGovern, P. G. and Folsom, A. R. (1992). Effect of family history, body-fat distribution, and reproductive factors on the risk of postmenopausal breast cancer. *N Engl J Med* **326**: 1323-9.
- Sellers, T. A., Gapstur, S. M., Potter, J. D., Kushi, L. H., Bostick, R. M. and Folsom, A. R. (1993). Association of body fat distribution and family histories of breast and ovarian cancer with risk of postmenopausal breast cancer. *Am J Epidemiol* **138**: 799-803.
- Seoud, M. A., Johnson, J. and Weed, J. J. (1993). Gynecologic tumors in tamoxifen-treated women with breast cancer. *Obstet Gynecol* **82**: 165-9.
- Serrano, M., Hannon, G. J. and Beach, D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* **366**: 704-7.

- Shamsher, M. K., Navsaria, H. A., Stevens, H. P., Ratnavel, R. C., Purkis, P. E., Kelsell, D. P., McLean, W. H. I., Cook, L. J., Griffiths, W. A. D., Gschmeissner, S., Spurr, N. and Leigh, I. M. (1995). Novel mutations in keratin-16 gene underlying focal non-epidermolytic palmoplantar keratoderma (NEPPK) in 2 families. *Hum Mol Genet* **4**: 1875-1881.
- Shattuck-Eidens, D., McClure, M., Simard, J., Labrie, F., Narod, S., Couch, F., Hoskins, K., Weber, B., Castilla, L., Erdos, M., Brody, L., *et al.* (1995). A collaborative survey of 80 mutations in the BRCA1 breast and ovarian cancer susceptibility gene: Implications for presymptomatic testing and screening. *J Am Med Assoc.* **273**: 535-541.
- Sherr, C.J. (1994) G1 phase progression-cycling on cue. *Cell.* **79**:551-555.
- Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**: 8794-7.
- Simard, J., Feunteun, J., Lenoir, G., Tonin, P., Normand, T., The, V. L., Vivier, A., Lasko, D., Morgan, K., Rouleau, G. A., Lynch, H., Labrie, F. and Narod, S. A. (1993). Genetic-mapping of the breast-ovarian cancer syndrome to a small interval on chromosome-17q12-21 - exclusion of candidate genes EDH17 β 2 and RARA. *Hum Mol Genet* **2**: 1193-1199.
- Sims, E. A. (1990). Destiny rides again as twins overeat. *N Engl J Med* **322**: 1522-4.
- Slamon, D. J., Clark, G. M., Wong, S. G. *et al.* (1987). Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235**: 177-182.
- Slamon, D. J., Godolphin, W., Jones, L. A., Holt, J. A., Wong, S. G., Keith, D. E., Levin, W. J., Stuart, S. G., Udove, J., Ullrich, A. and Press, M. F. (1989). Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* **244**: 707-712.
- Slattery, M. L. and Kerber, R. A. (1993). A comprehensive evaluation of family history and breast cancer risk. The Utah Population Database. *J Am Med Assoc.* **270**: 1563-8.
- Smith, S. A., Easton, D. F., Evans, D. and Ponder, B. (1992). Allele losses in the region 17q12-21 in familial breast and ovarian-cancer involve the wild-type chromosome. *Nat Genet* **2**: 128-131.
- Smith, S. A., Dicioccio, R. A., Struewing, J. P., Easton, D. F., Gallion, H. H., Albertsen, H., Mazoyer, S., Johansson, B., Steichen-Gersdorf, E. and Stratton, M. (1994). Localization of the breast-ovarian cancer susceptibility gene (BRCA1) on 17q12-21 to an interval of less than or equal to 1cM. *Genes Chromosomes Cancer* **10**: 71-76.
- Smith, S. A., Holik, P., Stevens, J., Mazoyer, S., Melis, R., Williams, B., White, R., Albertsen, H. (1996). Isolation of a gene (DLG3) encoding a second member of the Discs-Large family on chromosome 17q12-q21. *Genomics* **31**: 145-150.
- Solomon, E. and Ledbetter, D. H. (1990). Report of the committee on the genetic constitution of chromosome 17. *Cytogenet Cell Genet* **55**: 198-215.
- Solomon, E., Voss, R., Hall, V., Bodmer, W. F., Jass, J. R., Jeffreys, A. J., Lucibello, F. C., Patel, I. and Rider, S. H. (1987). Chromosome 5 allele loss in human colorectal carcinomas. *Nature* **328**: 616-9.
- Sorensen, P. H. B. and Triche, T. J. (1996). Gene fusions encoding chimaeric transcription factors in solid tumours. *Cancer Biol.* **7**:3-14.
- Southern, E. M. (1975). *J. Mol. Biol* **98**: 503.

- Sparkes, R. S., Murphree, A. L., Lingua, R. W., Sparkes, M. C., Field, L. L., Funderburk, S. J. and Benedict, W. F. (1983). Gene for hereditary retinoblastoma assigned to human chromosome 13 by linkage to esterase D. *Science* **219**: 971-3.
- Spurr, N. K., Kelsell, D., Rooke, L., Cavalli-Sforza, L. L., Bowcock, A. and Feder, J. (1991). Assignment of eight human genomic DNA sequences detecting high frequency restriction fragment length polymorphisms. *Ann Hum Genet* **55**: 141-50.
- Spurr, N. K., Kelsell, D. P., Black, D. M., Murday, V. A., Turner, G., Crockford, G. P., Solomon, E., Cartwright, R. A. and Bishop, D. T. (1993). Linkage analysis of early-onset breast and ovarian-cancer families, with markers on the long arm of chromosome-17. *Am J Hum Genet* **52**: 777-785.
- Spurr, N. K., Bouzyk, M., Kelsell, D. P. (1996). Gene mapping using somatic cell hybrids. *Methods in Molecular Biology*. (in press).
- Squire, J. and Weksberg, R. (1996). Genomic imprinting in tumours. *Cancer Biol.* **7**: 41-47.
- Srivastava, S., Zou, Z. Q., Pirollo, K., Blattner, W. and Chang, E. H. (1990). Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. *Nature* **348**: 747-9.
- Stahl, J. A., Leone, A., Rosengard, A. M., Porter, L., King, C. R. and Steeg, P. S. (1991). Identification of a second human nm23 gene, nm23-H2. *Cancer Res* **51**: 445-9.
- Steeg, P. S., Bevilacqua, G., Kopper, L., Thorgeirsson, U. P., Talmadge, J. E., Liotta, L. A. and Sobel, M. E. (1988). Evidence for a novel gene associated with low tumor metastatic potential. *J Natl Cancer Inst* **80**: 200-4.
- Steeg, P. S., and Liotta, L. A. (1990). Reduced nm23 expression in tumour metastasis. *Proc. Am. Assoc. Cancer. Res.* **31**: 504-505.
- Steichen-Gersdorf, E., Gallion, H. H., Ford, D., Girodet, C., Easton, D. F., DiCioccio, R. A., Evans, G., Ponder, M. A., Pye, C., Mazoyer, S., et al. (1994). Familial site-specific ovarian cancer is linked to BRCA1 on 17q12-21. *Am J Hum Genet* **55**: 870-875.
- Steinert, P. M. (1993). Structure, function, and dynamics of keratin intermediate filaments. *J Invest Dermatol* **100**: 729-34.
- Stevens, H. P., Kelsell, D. P., Bryant, S. P., Spurr, N. K., Weissenbach, J., Marger, D., Marger, R. S., Leigh, I. M. (1996a). Palmoplantar keratoderma and malignancy (palmoplantar keratoderma ectodermal dysplasia type III) in an extensive North American pedigree is linked to the TOC locus (17q24) and show a characteristic clinical phenotype. With bibliography and proposed classification of the palmoplantar keratodermas. *Arch. Dermatol.* **132**: 640-651.
- Stevens, H. P., Kelsell, D. P., Leigh, I. M., Ostlere, L. S., MacDermot, K. D., Rustin, M. J. A. (1996b). Punctate palmoplantar keratoderma and malignancy in a four-generation family. *Br. J. Dermatol.* **134**: 720-726.
- Stoll, B. A. (1994). Breast-cancer - the obesity connection. *Br J Cancer* **69**: 799-801.
- Strachan, T. S. and Read, A. P. (1996). *Human Molecular Genetics*. BIOS Scientific Publishers Limited.
- Stratton, M. R., Ford, D., Neuhasen, S., Seal, S., Wooster, R., Friedman, L. S., King, M. C., Egilsson, V., Devilee, P. and McManus, R. (1994). Familial male breast-cancer is not linked to the BRCA1 locus on chromosome 17q. *Nature Genet* **7**: 103-107.

Struewing, J. P., Abeliovich, D., Peretz, T., Avishai, N., Kaback, M. M., Collins, F. S. and Brody, L. C. (1995a). The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi-Jewish individuals. *Nature Genet* 11: 198-200.

Struewing, J. P., Brody, L. C., Erdos, M. R., Kase, R. G., Giambarresi, T. R., Smith, S. A., Collins, F. S. and Tucker, M. A. (1995b). Detection of 8 BRCA1 mutations in 10 breast ovarian-cancer families, including 1 family with male breast-cancer. *Am J Hum Genet* 57: 1-7.

Su, L. K., Vogelstein, B. and Kinzler, K. W. (1993). Association of the APC tumor suppressor protein with catenins. *Science* 262: 1734-7.

Sutherland, R. L., Watts, C. K. and Musgrove, E. A. (1993). Cyclin gene expression and growth control in normal and neoplastic human breast epithelium. *J Steroid Biochem Mol Biol* 47: 99-106.

Swift, M., Morrell, D., Massey, R. B. and Chase, C. L. (1992). Risk of breast-cancer in ataxia telangiectasia. *N Engl J Med* 326: 1359-1360.

Szabo, C. I. and King, M. C. (1995). Inherited breast and ovarian-cancer. *Hum Mol Genet* 4: 1811-1817.

T'Ang, A., Varley, J. M., Chakraborty, S., Murphree, A. L. and Fung, Y. K. (1988). Structural rearrangement of the retinoblastoma gene in human breast carcinoma. *Science* 242: 263-6.

Tanner, M. M., Tirkkonen, M., Kallioniemi, A., Collins, C., Stokke, T., Karhu, R., Kowbel, D., Shadravan, F., Hintz, M., Kuo, W. L., *et al.* (1994). Increased copy number at 20q13 in breast cancer: defining the critical region and exclusion of candidate genes. *Cancer Res* 54: 4257-60.

Tavtigian, S. V., Simard, J., Rommens, J., Couch, F., Shattuck-Eidens, D., Neuhausen, S., Merajver, S., Thorlacius, S., Offit, K., Stoppa, L. D., *et al.* (1996). The complete BRCA2 gene and mutations in chromosome 13q-linked kindreds. *Nat Genet* 12: 333-7.

Teng, D. H-F., Bogden, R., Mitchell, J., Baumgard, M., Bell, R., Berry, S., Davis, T., Ha, P-C., Kehrer, R., *et al.* (1996). Low incidence of BRCA2 mutations in breast carcinoma and other cancers. *Nat Genet*. 13: 241-244.

The, V. L., Labrie, C., Zhao, H. F., Couet, J., Lachance, Y., Simard, J., Leblanc, G., Cote, J., Berube, D., Gagne, R. *et al.* (1989). Characterization of cDNAs for human estradiol 17 beta-dehydrogenase and assignment of the gene to chromosome 17: evidence of two mRNA species with distinct 5'-termini in human placenta. *Mol Endocrinol* 3: 1301-9.

Thompson, M. E., Jensen, R. A., Obermiller, P. S., Page, D. L. and Holt, J. T. (1995). Decreased expression of BRCA1 accelerates growth and is often present during sporadic breast cancer progression. *Nat Genet* 9: 444-50.

Thorlacius, S., Tryggvadottir, L., Olafsdottir, G. H., Jonasson, J. G., Ogmundsdottir, H. M., Tulinius, H. and Eyfjord, J. E. (1995). Linkage to BRCA2 region in hereditary male breast-cancer. *Lancet* 346: 544-545.

Thorlacius, S., Olafsdottir, G., Tryggvadottir, L., Neuhausen, S., Jonasson, J. G., Tavtigian, S. V., Tulinius, H., Ogmundsdottir, H. M., Eyfjord, J. E. (1996). A single BRCA2 mutation in male and female breast cancer families from Iceland with varied cancer phenotypes. *Nat. Genet*. 13: 117-119.

Tonin, P., Ghadirian, P., Phelan, C., Lenoir, G. M., Lynch, H. T., Letendre, F., Belanger, D., Monte, M., Narod, S. A. (1995). A large multisite cancer family is linked to BRCA2. *J. Med. Genet* 32: 982-984.

- Torchard, D., Blanchet-Bardon, C., Serova, O., Langbein, L., Narod, S., Janin, N., Goguel, A. F., Bernheim, A., Franke, W. W., Lenoir, G. M., *et al.* (1994). Epidermolytic palmoplantar keratoderma cosegregates with a keratin 9 mutation in a pedigree with breast and ovarian cancer. *Nat Genet* **6**: 106-10.
- Trofatter, J. A., MacCollin, M. M., Rutter, J. L., Murrell, J. R., Duyao, M. P., Parry, D. M., Eldridge, R., Kley, N., Menon, A. G., Pulaski, K., *et al.* (1993). A novel moesin-ezrin-radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. *Cell* **72**: 791-800.
- Tulinius, H., Egilsson, V., Olafsdottir, G. H. and Sigvaldason, H. (1992). Risk of prostate, ovarian, and endometrial cancer among relatives of women with breast-cancer. *BMJ*. **305**: 855-857.
- Varesco, L., Caligo, M. A., Simi, P., Black, D. M., Nardini, V., Casarino, L., Rocchi, M., Ferrara, G., Solomon, E. and Bevilacqua, G. (1992). The NM23 gene maps to human chromosome band 17q22 and shows a restriction fragment length polymorphism with BglII. *Genes Chromosomes Cancer* **4**: 84-8.
- Varley, J. M., Brammar, W. J., Lane, D. P., Swallow, J. E., Dolan, C. and Walker, R. A. (1991). Loss of chromosome 17p13 sequences and mutation of p53 in human breast carcinomas. *Oncogene* **6**: 413-21.
- Velu, T. J., Beguinot, L., Vass, W. C., Willingham, M. C., Merlino, G. T., Pastan, I. and Lowy, D. R. (1987). Epidermal-growth-factor-dependent transformation by a human EGF receptor proto-oncogene. *Science* **238**: 1408-10.
- Vermeulen, A., Deslypere, J. P., Paridaens, R., Leclercq, G., Roy, F. and Heuson, J. C. (1986). Aromatase, 17 beta-hydroxysteroid dehydrogenase and intratissular sex hormone concentrations in cancerous and normal glandular breast tissue in postmenopausal women. *Eur J Cancer Clin Oncol* **22**: 515-25.
- Vogelstein, B., Fearon, E. R., Kern, S. E., Hamilton, S. R., Preisinger, A. C., Nakamura, Y. and White, R. (1989). Allelotype of colorectal carcinomas. *Science* **244**: 207-11.
- Vogelstein, B. and Kinzler, K. W. (1993). The multistep nature of cancer. *Trends Genet* **9**: 138-141.
- Wales, M. M., Biel, M. A., el Deiry, W., Nelkin, B. D., Issa, J. P., Cavenee, W. K., Kuerbitz, S. J. and Baylin, S. B. (1995). p53 activates expression of HIC-1, a new candidate tumour suppressor gene on 17p13.3. *Nat Med* **1**: 570-7.
- Wallet, V., Mutzel, R., Troll, H., Barzu, O., Wurster, B., Veron, M. and Lacombe, M. L. (1990). Dictyostelium nucleoside diphosphate kinase highly homologous to Nm23 and Awd proteins involved in mammalian tumor metastasis and Drosophila development. *J Natl Cancer Inst* **82**: 1199-202.
- Wang, T. C., Cardiff, R. D., Zukerberg, L., Lees, E., Arnold, A. and Schmidt, E. V. (1994). Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. *Nature* **369**: 669-71.
- Washington, S. S., Warburton, D. and Chakravarti, A. (1995). Report and abstracts of the Second International Workshop on Human Chromosome 13 Mapping 1994. *Cytogenet Cell Genet* **70**: 1-22.
- Watkins, C., Warne, D., Kelsell, D., Beckmann, J., Nyberg, K., Spurr, N. (1991). The use of PCR and ethidium bromide staining to detect length variation in short repeat units. *Technique* **3**: 175-178.

- Weber, J. L. and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* **44**: 388-96.
- Weber, J. L. (1990). Informativeness of human (dC-dA)_n.(dG-dT)_n polymorphisms. *Genomics* **7**: 524-30.
- Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. and Lathrop, M. (1992). A 2nd-generation linkage map of the human genome. *Nature* **359**: 794-801.
- Weissman, B. E. and Stanbridge, E. J. (1983). Complementation of the tumorigenic phenotype in human cell hybrids. *J Natl Cancer Inst* **70**: 667-72.
- White, R. and Lalouel, J. M. (1988). Chromosome mapping with DNA markers. *Sci Am* **258**: 40-8.
- Willard, H. F., Skolnick, M. H., Pearson, P. L. and Mandel, J. L. (1985). Report of the Committee on Human Gene Mapping by Recombinant DNA Techniques. *Cytogenet Cell Genet* **40**: 360-489.
- Woods, D. F. and Bryant, P. J. (1989). Molecular cloning of the lethal(1)discs large-1 oncogene of *Drosophila*. *Dev Biol* **134**: 222-35.
- Wooster, R., Mangion, J., Eeles, R., Smith, S., Dowsett, M., Averill, D., Barretlee, P., Easton, D. F., Ponder, B. and Stratton, M. R. (1992). A germline mutation in the androgen receptor gene in 2 brothers with breast-cancer and Reifenstein syndrome. *Nat Genet* **2**: 132-134.
- Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., *et al.* (1994). Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. *Science* **265**: 2088-90.
- Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., *et al.* (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**: 789-92.
- Xiong, Y., Hannon, G. J., Zhang, H., Casso, D., Kobayashi, R. and Beach, D. (1993). p21 is a universal inhibitor of cyclin kinases. *Nature* **366**: 701-4.
- Xu, C-F., Chambers, J.A., Nicolai, H., Brown, M.A., Hujeirat, Y., Mohammed, S., Hodgeson, S., Kelsell, D.P., Spurr, N.K., Bishop, D.T., Solomon, E. (1996). Mutations and alternative splicing of the BRCA1 gene in UK breast-ovarian cancer families. *Genes, Chromosomes and Cancer*. (in press).
- Zhang, Y., Shields, T., Crenshaw, T., Hao, Y., Moulton, T. and Tycko, B. (1993). Imprinting of human H19: allele-specific CpG methylation, loss of the active allele in Wilms tumor, and potential for somatic allele switching. *Am J Hum Genet* **53**: 113-24.
- Ziegle, J. S., Su, Y., Corcoran, K. P., Nie, L., Mayrand, P. E., Hoff, L. B., McBride, L. J., Kronick, M. N. and Diehl, S. R. (1992). Application of automated DNA sizing technology for genotyping microsatellite loci. *Genomics* **14**: 1026-1031.
- Zuo, L., Weger, J., Yang, Q., Goldstein, A. M., Tucker, M. A., Walker, G. J., Hayward, N. and Dracopoli, N. C. (1996). Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nat Genet* **12**: 97-9.